Balancing Fairness and Efficiency in an Optimization Model

John Hooker
Carnegie Mellon University

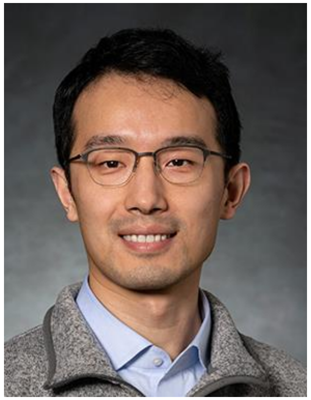
University of Toronto
November 2021

Much of this work is joint with:

Violet (Xinying) Chen, *CMU*

Özgün Elçi
CMU

H. P. Williams
London School of Economics



Peter Zhang
CMU

Modeling Equity

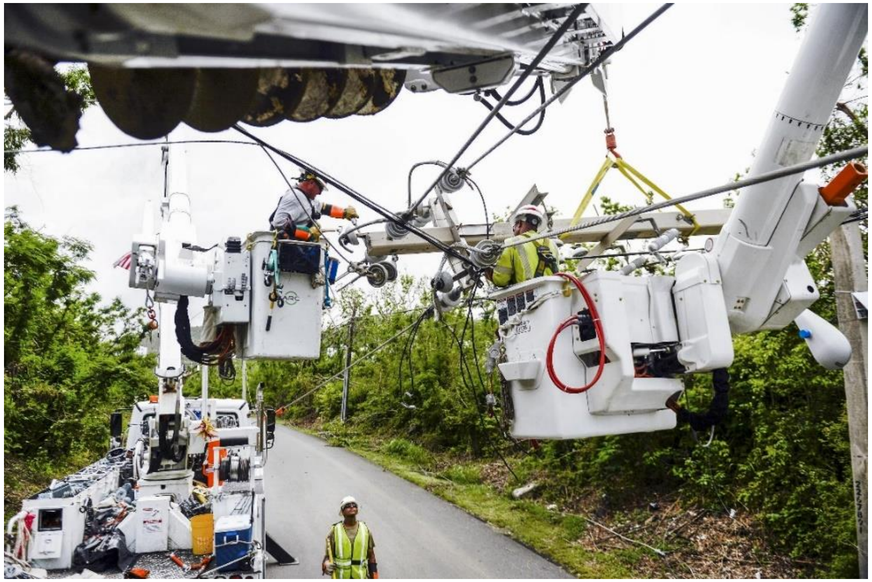
- A growing interest in incorporating **equity** into models…
    - Health care resources.
    - Facility location (e.g., emergency services, infrastructure).
    - Taxation (revenue vs. progressivity).
    - Telecommunications (leximax, Nash bargaining solution).
    - Traffic signal timing
    - Disaster recovery (e.g., power restoration)…

UNHCR
The UN Refugee Agency
Ύπατη Αρμοστεία του ΟΗΕ
για τους Πρόσφυγες

Modeling Equity

- Example: disaster relief
  - Power restoration can focus on **urban** areas first (**efficiency**).
  - This can leave rural areas without power for weeks/months.
  - This happened in Puerto Rico after Hurricane Maria (2017).

– A more **equitable** solution
  - …would give some priority to rural areas without overly sacrificing efficiency.



Modeling Equity

- It is far from obvious how to formulate equity concerns **mathematically**.
    - Less straightforward than maximizing total benefit or minimizing total cost.
    - Still less obvious how to **combine** equity with total benefit.

Outline

- Generic welfare optimization model
- Existing formulations of equity
  - **Inequality**-based criteria
  - Fairness for the **disadvantaged** (Rawlsian maximin, McLoone)
  - **Convex combinations** of utility and equity
  - **Alpha fairness** and **Nash bargaining** solution
  - **Kalai-Smorodinsky** bargaining solution
  - **Statistical fairness** metrics used in AI
  - **Utility-** and **equity-threshold** criteria combining utility & maximin
- Our most recent proposal
  - **Utility-threshold** criterion combining utility and **leximax**
  - Examples: **health care** and **earthquake shelter location**

Generic Model

- We formulate each fairness criterion as a **social welfare function (SWF)**.

Individual utilities

$$W(\boldsymbol{u}) = W(u_1, \ldots, u_n)$$

  - Measures desirability of the **magnitude and distribution of utilities** across individuals.
  - **Utility** can be wealth, health, negative cost, etc.

Generic Model

The social welfare optimization problem

Vector of utilities enjoyed by individuals

Vector of resource allocations $(x_1, \ldots, x_n)$ to individuals

$$\max_{\boldsymbol{u},\boldsymbol{x}} \left\{ W(\boldsymbol{u}) \;\middle|\; (\boldsymbol{u}, \boldsymbol{x}) \in S \right\}$$

Social welfare function

Problem constraints, including constraints that define $\boldsymbol{u}$ in terms of other variables $\boldsymbol{x}$

Generic Model

Example – *Medical triage*

QALYs without treatment

Additional QALYs due to treatment

= 1 if patient *i* is treated

Utility functions are $U_i(\boldsymbol{x}) = a_i + b_i x_i$.

$$\max_{\boldsymbol{u},\boldsymbol{x}} \left\{ W(\boldsymbol{u}) \;\middle|\; \begin{array}{l} \sum_i c_i x_i \le B \\ u_i = a_i + b_i x_i,\; x_i \in \{0,1\},\; \text{all } i \end{array} \right\}$$

Social welfare function

Budget constraint

Yes-or-no decision

Generic Model

What we want to contribute to practice:

Show how to add equity considerations to an existing optimization model.

- **Utility** is already defined in the model.
- Identify a suitable **social welfare function** that can serve as the **objective function** of the model.

Inequality Measures

Equality vs fairness

Two views on ethical importance of equality: Parfit 1997

- **Irreducible:** Inequality is inherently unfair. Scanlon 2003
- **Reducible:** Inequality is unfair only insofar as it reduces utility.

Frankfurt 2015

Problems with inequality measures:

- No preference for an identical distribution with **higher utility**.
- Even when average utility is fixed, no preference for reducing inequality at the **bottom** rather than the **top** of the distribution.

Inequality Measures

Relative range

$$W(\boldsymbol{u}) = -\frac{u_{\text{max}} - u_{\text{min}}}{\bar{u}}$$

Rationale:

- Perceived inequality is relative to the best off.
- So, move everyone closer to the best off.

Problem:

- Ignores distribution **between** extremes.

Inequality Measures

Relative range

- Problem is **linearized** using same change of variable as in linear-fractional programming.

Let $\boldsymbol{u} = \boldsymbol{u}'/t$ and $\boldsymbol{x} = \boldsymbol{x}'/t$. The optimization problem is

$$\min_{\substack{\boldsymbol{x}',\boldsymbol{u}',t \\ u'_{\min},u'_{\max}}} \left\{ u'_{\max} - u'_{\min} \;\middle|\; \begin{array}{c} u'_{\min} \le u'_i \le u'_{\max},\ \text{all } i \\ \bar{u}' = 1,\ t \ge 0,\ (\boldsymbol{u}', \boldsymbol{x}') \in S' \end{array} \right\}$$

where $t, u'_{\min}, u'_{\max}$ are new variables.

Charnes & Cooper 1962

- Linear if original constraints $(\boldsymbol{u}, \boldsymbol{x}) \in S$ are linear.

Inequality Measures

Relative mean deviation

$$W(\boldsymbol{u}) = -\frac{1}{\bar{u}} \sum_i |u_i - \bar{u}|$$

Rationale:

- Considers all utilities.

Model:

- Again, linearized by change of variable.

$$\min_{\boldsymbol{x}', \boldsymbol{u}', \boldsymbol{v}, t} \left\{ \sum_i v_i \;\middle|\; \begin{array}{c} -v_i \leq u_i' - \bar{u}' \leq v_i, \text{ all } i \\ \bar{u}' = 1, \; t \geq 0, \; (\boldsymbol{u}', \boldsymbol{x}') \in S' \end{array} \right\}$$

where $\boldsymbol{v}$ is vector of new variables.

Inequality Measures

Coefficient of variation

$$W(\boldsymbol{u}) = -\frac{1}{\bar{u}}\left[\frac{1}{n}\sum_i (u_i - \bar{u})^2\right]^{\frac{1}{2}}$$

Rationale:

- Familiar. Outliers receive extra weight.

Problem:

- Nonlinear (but convex)

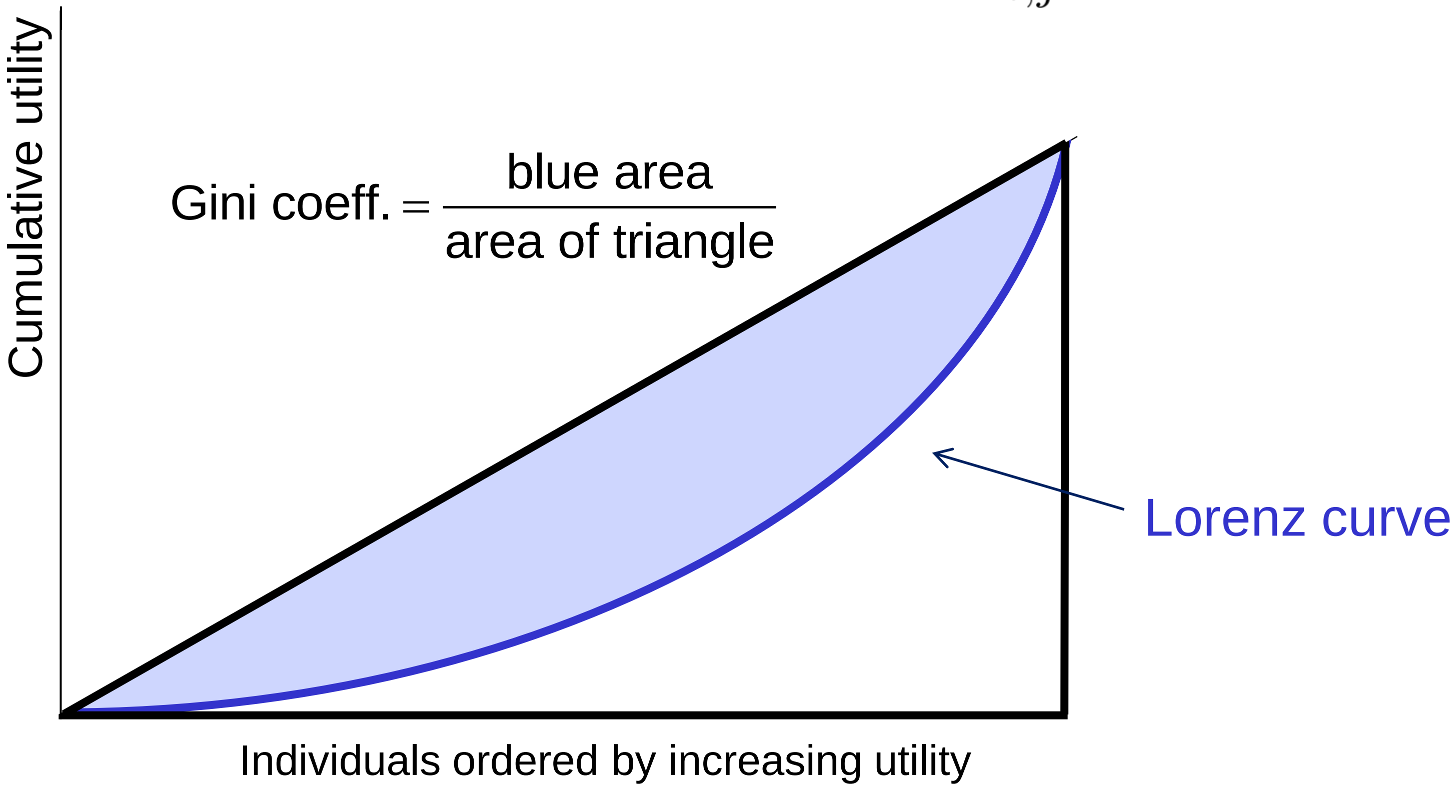
Model:

$$\min_{\boldsymbol{x}',\boldsymbol{u}',\boldsymbol{v},t} \left\{ \frac{1}{n}\sum_i (u_i' - \bar{u}')^2 \;\middle|\; \begin{array}{c} \bar{u}' = 1,\; t \geq 0 \\ (\boldsymbol{u}', \boldsymbol{x}') \in S' \end{array} \right\}$$

Inequality Measures

Gini coefficient

$$W(\boldsymbol{u}) = -G(\boldsymbol{u}), \quad \text{where } G(\boldsymbol{u}) = \frac{1}{2\bar{u}n^2} \sum_{i,j} |u_i - u_j|$$



Inequality Measures

Gini coefficient

$$W(\boldsymbol{u}) = -G(\boldsymbol{u}), \quad \text{where } G(\boldsymbol{u}) = \frac{1}{2\bar{u}n^2} \sum_{i,j} |u_i - u_j|$$

Rationale:

- Relationship to Lorenz curve.
- Widely used.

Model:

- Linear:

$$\min_{\boldsymbol{x}', \boldsymbol{u}', V, t} \left\{ \frac{1}{2n^2} \sum_{i,j} v_{ij} \;\middle|\; \begin{array}{c} -v_{ij} \le u'_i - u'_j \le v_{ij}, \text{ all } i,j \\ \bar{u}' = 1,\ t \ge 0,\ (\boldsymbol{u}', \boldsymbol{x}') \in S' \end{array} \right\}$$

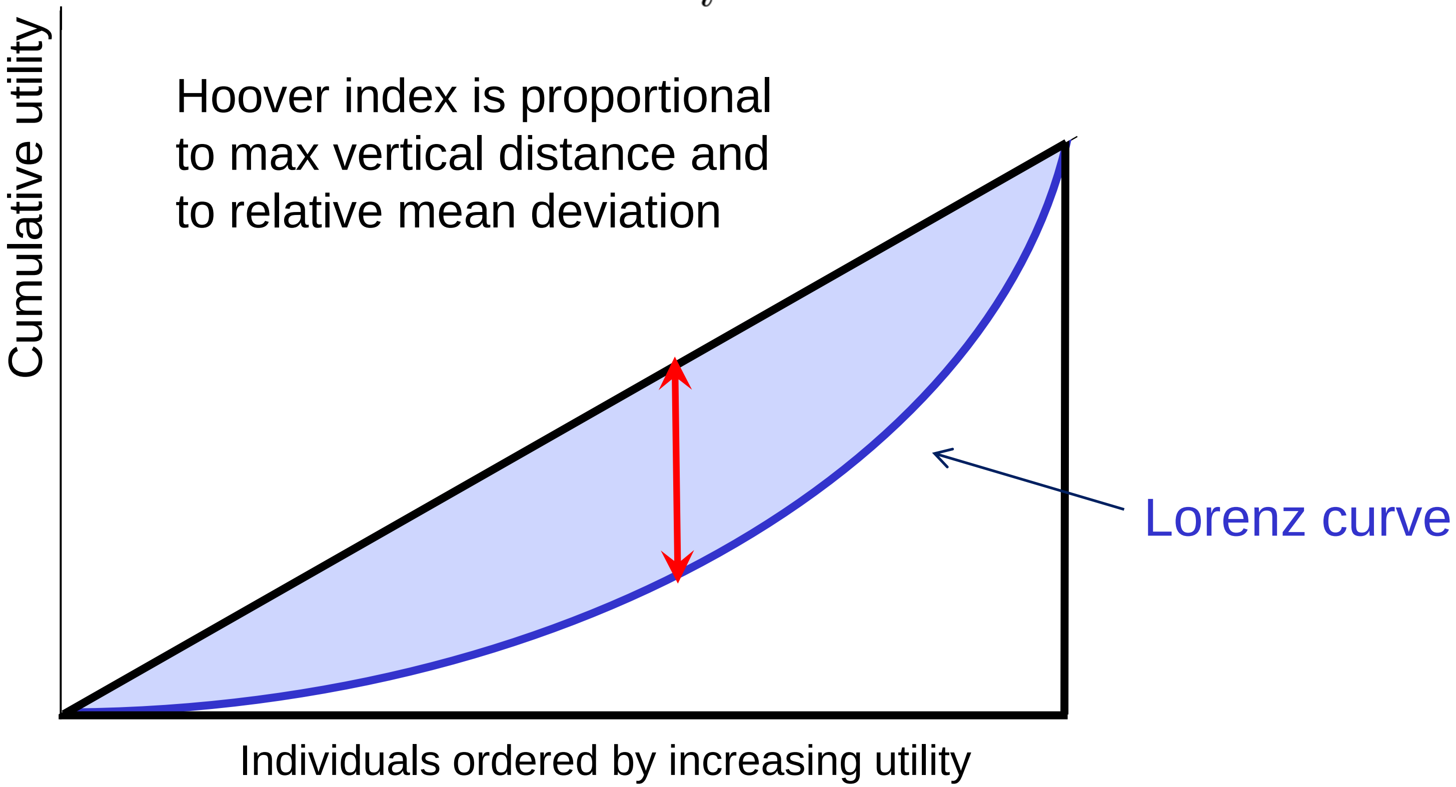
where V is a matrix of new variables.

Inequality Measures

Hoover index

$$W(\boldsymbol{u}) = -\frac{1}{2n\bar{u}} \sum_i |u_i - \bar{u}|$$

Hoover 1936



Inequality Measures

Hoover index

$$W(\boldsymbol{u}) = -\frac{1}{2n\bar{u}} \sum_i |u_i - \bar{u}|$$

Rationale:

- Hoover index is fraction of total utility that would have to be redistributed to achieve perfect equality.

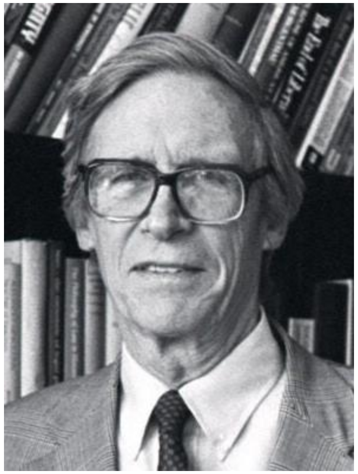
Model:

- Same as relative mean deviation.

Fairness for the Disadvantaged

Maximin

$$W(\boldsymbol{u}) = \min_i \{u_i\}$$



Rationale:

- Based on **difference principle** of John Rawls.
- Inequality is justified only to the extent that it increases the utility of the worst-off.
- Originally intended only for the design of **social institutions** and distribution of **primary goods** (goods that any rational person would want).
- Can be adopted as a general principle of equity: maximize the minimum utility.

Rawls 1971, 1999

Fairness for the Disadvantaged

Maximin

$$W(\boldsymbol{u}) = \min_i \{u_i\}$$

Social contract argument:

- We decide on social policy in an "original position," behind a "veil of ignorance" as to our position on society.
- All parties must be willing to **endorse** the policy, no matter what position they end up assuming.
- No rational person can endorse a policy that puts him/her on the **bottom** of society – unless that person would be even **worse off** under another social arrangement.
- Therefore, an agreed-upon social policy must maximize the welfare of the worst-off.

Fairness for the Disadvantaged

Maximin

$$W(\boldsymbol{u}) = \min_i \{u_i\}$$

Model:

$$\max_{\boldsymbol{x},\boldsymbol{u},w} \left\{ w \;\middle|\; w \leq u_i,\ \text{all } i;\ (\boldsymbol{u},\boldsymbol{x}) \in S \right\}$$

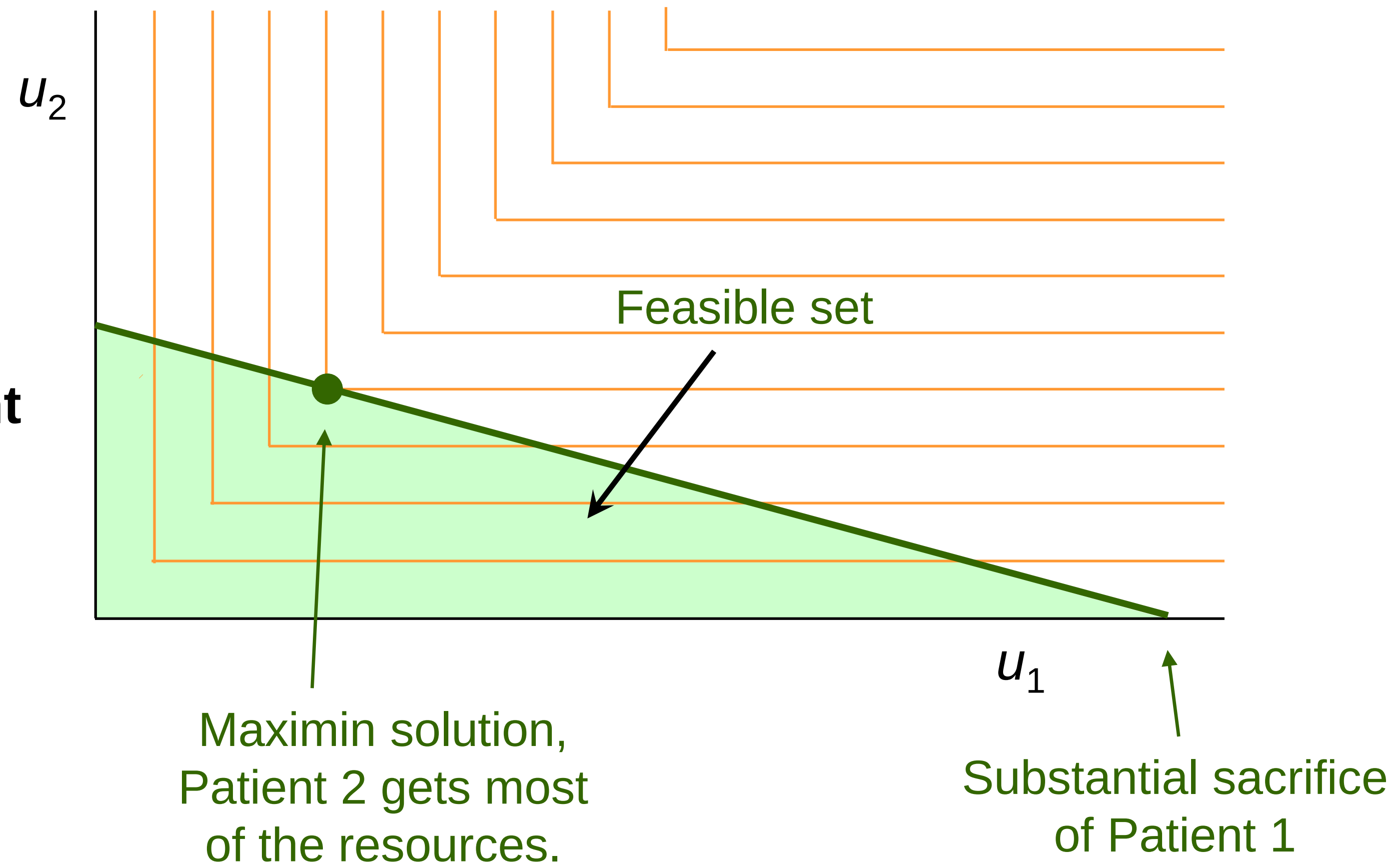
Problems:

- Can force equality even when this is extremely costly in terms of total utility.
- Does not care about 2$^{\text{nd}}$ worst off, etc., and so can waste resources.

Fairness for the Disadvantaged

Maximin

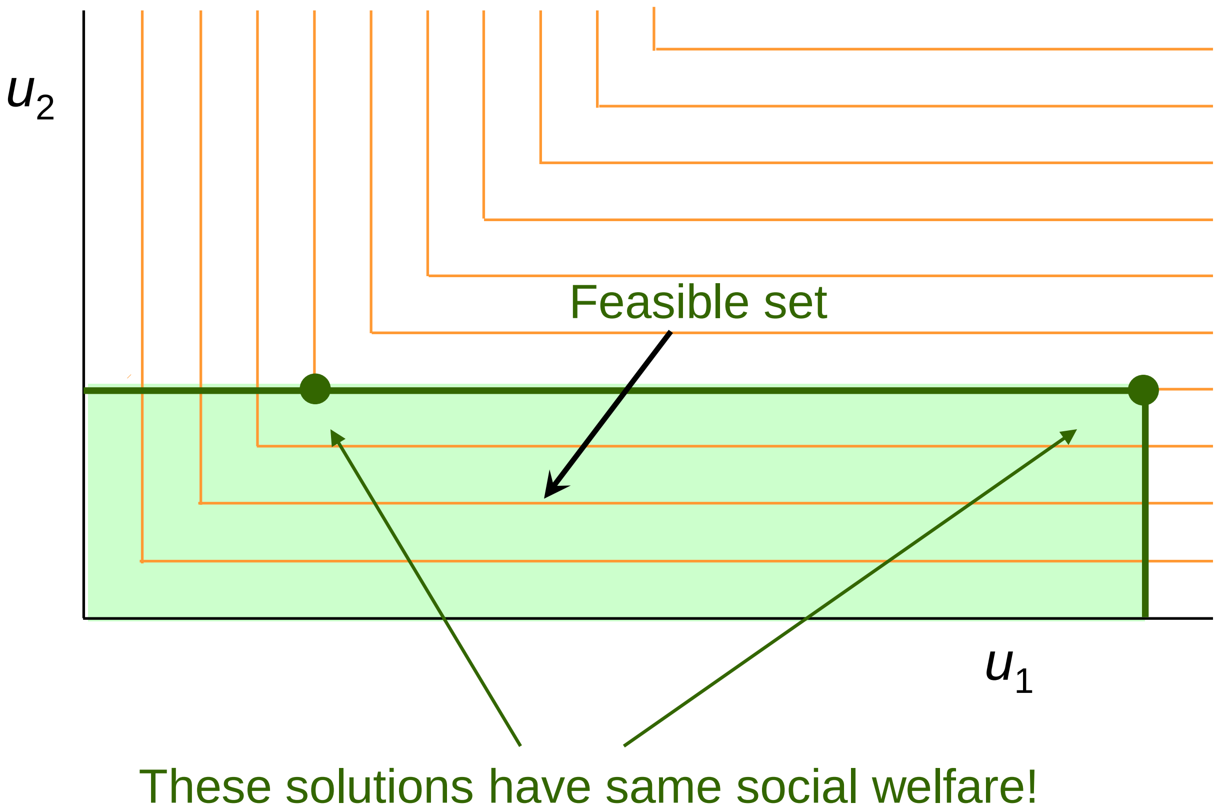
Medical example with budget constraint



Fairness for the Disadvantaged

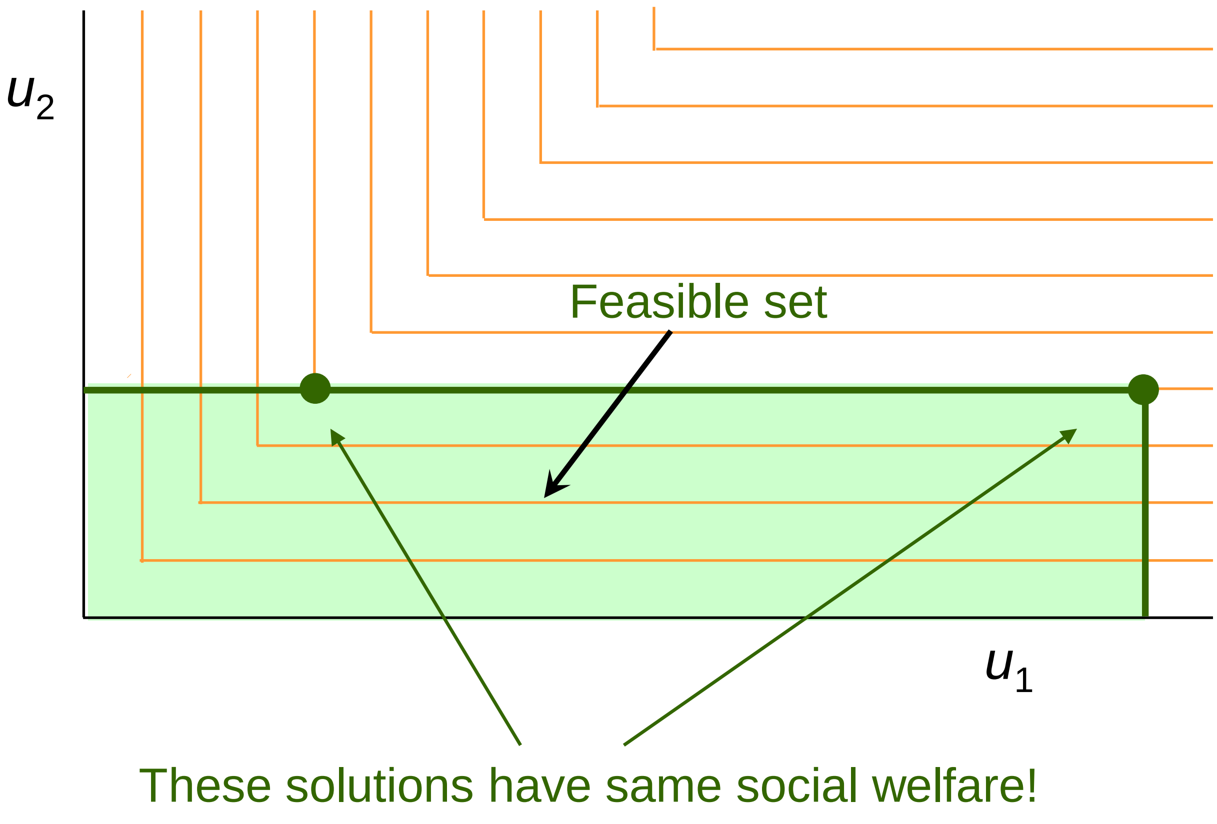
Maximin

Medical example with resource bounds



Fairness for the Disadvantaged

Maximin



Medical example with resource bounds

Remedy**:** use **leximax** solution

Fairness for the Disadvantaged

Leximax

Rationale:

- Takes in account 2nd worst-off, etc., and avoids wasting utility.
- Can be justified with Rawlsian argument.

Problem:

- No practical SWF for leximax.
- Must solve sequence of max problems.
- Even this requires enumeration of all ties to ensure that leximax is found.

Fairness for the Disadvantaged

McLoone index

$$W(\boldsymbol{u}) = \frac{1}{|I(\boldsymbol{u})|\tilde{u}} \sum_{i \in I(\boldsymbol{u})} u_i$$

where $\tilde{u}$ is the median of utilities in $\boldsymbol{u}$ and $I(\boldsymbol{u})$ is the set of indices of utilities at or below the median

Rationale:

- Compares total utility of those at or below the median to the utility that would result from bringing them up to the median.
- Index = 1 if no one is below median, $\rightarrow$ 0 for long lower tail.
- Focus on **all** the **disadvantaged**.
- Often used for public goods (e.g., educational benefits).

Fairness for the Disadvantaged

McLoone index

Model: Nonlinear, requires 0-1 variables.

$$\max_{\substack{\boldsymbol{x},\boldsymbol{u},m\\ \boldsymbol{y},\boldsymbol{z},\boldsymbol{\delta}}} \left\{ \frac{\sum_i y_i}{\sum_i z_i} \;\middle|\; \begin{array}{c} m - M\delta_i \le u_i \le m + M(1-\delta_i),\ \text{all } i \\ y_i \le u_i,\ y_i \le M\delta_i,\ \delta_i \in \{0,1\},\ \text{all } i \\ z_i \ge 0,\ z_i \ge m - M(1-\delta_i),\ \text{all } i \\ \sum_i \delta_i \le n/2,\ (\boldsymbol{u},\boldsymbol{x}) \in S \end{array} \right\}$$

Linearize with change of variable, obtain MILP.

$$\max_{\substack{\boldsymbol{x}',\boldsymbol{u}',m'\\ \boldsymbol{y}',\boldsymbol{z}',t,\boldsymbol{\delta}}} \left\{ \sum_i y'_i \;\middle|\; \begin{array}{c} u'_i \ge m' - M\delta_i, \text{all } i \\ u'_i \le m' + M(1-\delta_i),\ \text{all } i \\ y'_i \le u'_i,\ y'_i \le M\delta_i,\ \delta_i \in \{0,1\},\ \text{all } i \\ z'_i \ge 0,\ z'_i \ge m' - M(1-\delta_i),\ \text{all } i \\ \sum_i z'_i = 1,\ t \ge 0 \\ \sum_i \delta_i \le n/2,\ (\boldsymbol{u}',\boldsymbol{x}') \in S' \end{array} \right\}$$

Utility & Fairness – Convex Combinations

Utility + Gini coefficient

$$W(\boldsymbol{u}) = (1-\lambda)\sum_i u_i + \lambda\big(1 - G(\boldsymbol{u})\big)$$

Rationale.

- Takes into account both efficiency and equity.
- Allows one to adjust their relative importance.

Problem.

- Combines utility with a dimensionless quantity.
- Choice of λ is an issue with convex combinations in general.

Utility & Fairness – Convex Combinations

Utility + Maximin

$$W(\boldsymbol{u}) = (1-\lambda)\sum_i u_i + \lambda \min_i\{u_i\}$$

Rationale.

- Explicitly considers individuals other than worst off.

Problem.

- If u_k is smallest utility, this is simply the linear combination

$$W(\boldsymbol{u}) = u_k + (1-\lambda)\sum_{i \neq k} u_i$$

- How to interpret λ?

Alpha Fairness

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

Mo & Walrand 2000; Verloop, Ayesta & Borst 2010

Rationale.

- Continuous and well-defined adjustment of equity/efficiency tradeoff.

  Utility u_j must be reduced by $(u_j/u_i)^\alpha$ units to compensate for a unit increase in u_i $(< u_j)$ while maintaining constant social welfare.

- Integral of power law $\Sigma_i u_i^{-\alpha}$
- Utilitarian when $\alpha = 0$, maximin when $\alpha \to \infty$

Alpha Fairness

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

Model

- Nonlinear but concave.

$$\max_{\boldsymbol{x},\boldsymbol{u}} \left\{ W_\alpha(\boldsymbol{u}) \;\middle|\; (\boldsymbol{u},\boldsymbol{x}) \in S \right\}$$

- Can be solved by efficient algorithms if constraints are linear (or perhaps if *S* is convex).

Alpha Fairness

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

Problems

- Unclear how to choose α in practice.
- An egalitarian distribution can have same social welfare as arbitrarily extreme inequality.

In a 2-person problem, the distribution $(u_1, u_2) = (1, 1)$ has the same social welfare as $(2^{1/(1-\alpha)}, \infty)$ when $\alpha > 1$.

Proportional Fairness

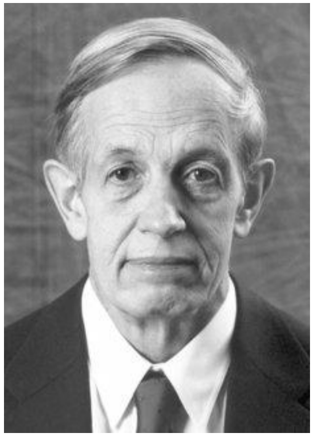
$$W(\boldsymbol{u}) = \sum_i \log(u_i)$$

Nash 1950

- Special case of alpha fairness (α = 1).
- Also known as **Nash bargaining solution**, in which case bargaining starts with a default distribution ***d***.

$$W(\boldsymbol{u}) = \sum_i \log(u_i - d_i) \quad \text{or}$$

$$W(\boldsymbol{u}) = \prod_i (u_i - d_i)$$



Proportional Fairness

$$W(\boldsymbol{u}) = \sum_i \log(u_i)$$

Nash 1950

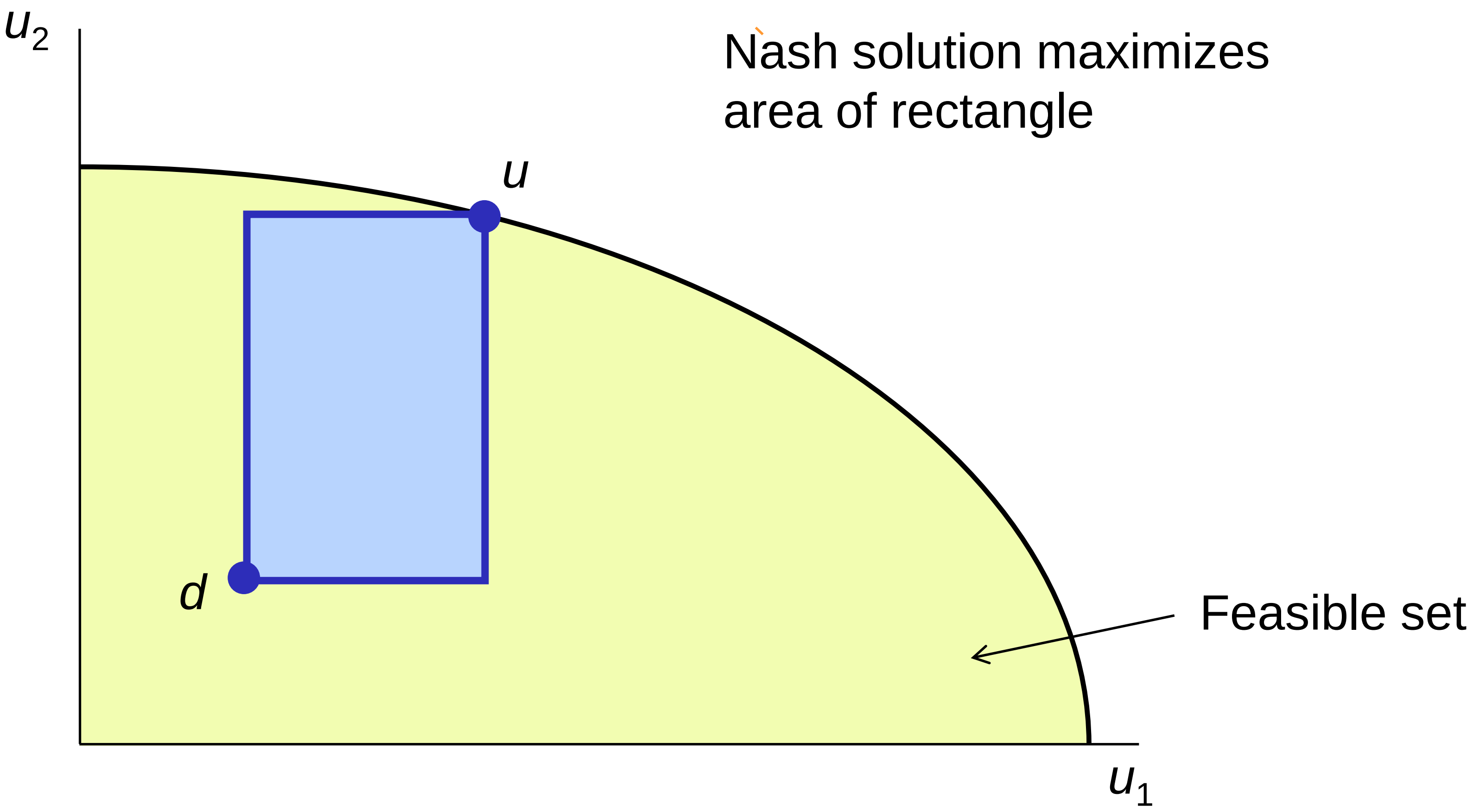
- Special case of alpha fairness (α = 1).
- Also known as **Nash bargaining solution**, in which case bargaining starts with a default distribution ***d***.

$$W(\boldsymbol{u}) = \sum_i \log(u_i - d_i) \quad \text{or} \quad W(\boldsymbol{u}) = \prod_i (u_i - d_i)$$

Rationale

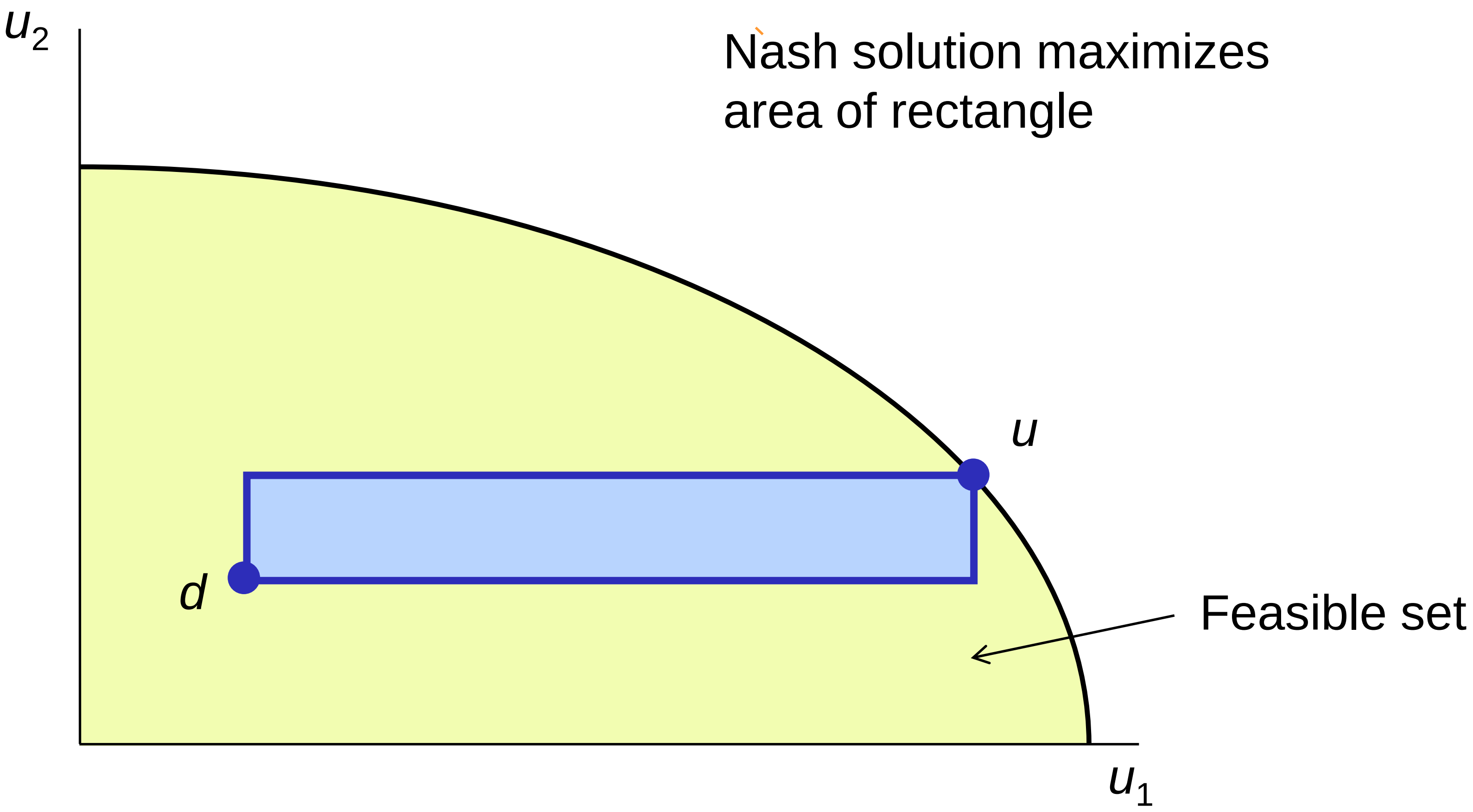
- Has nice geometric interpretation.
- Can be derived from axiomatic and bargaining arguments.
- Used in engineering applications (telecom, traffic signaling).

Proportional Fairness

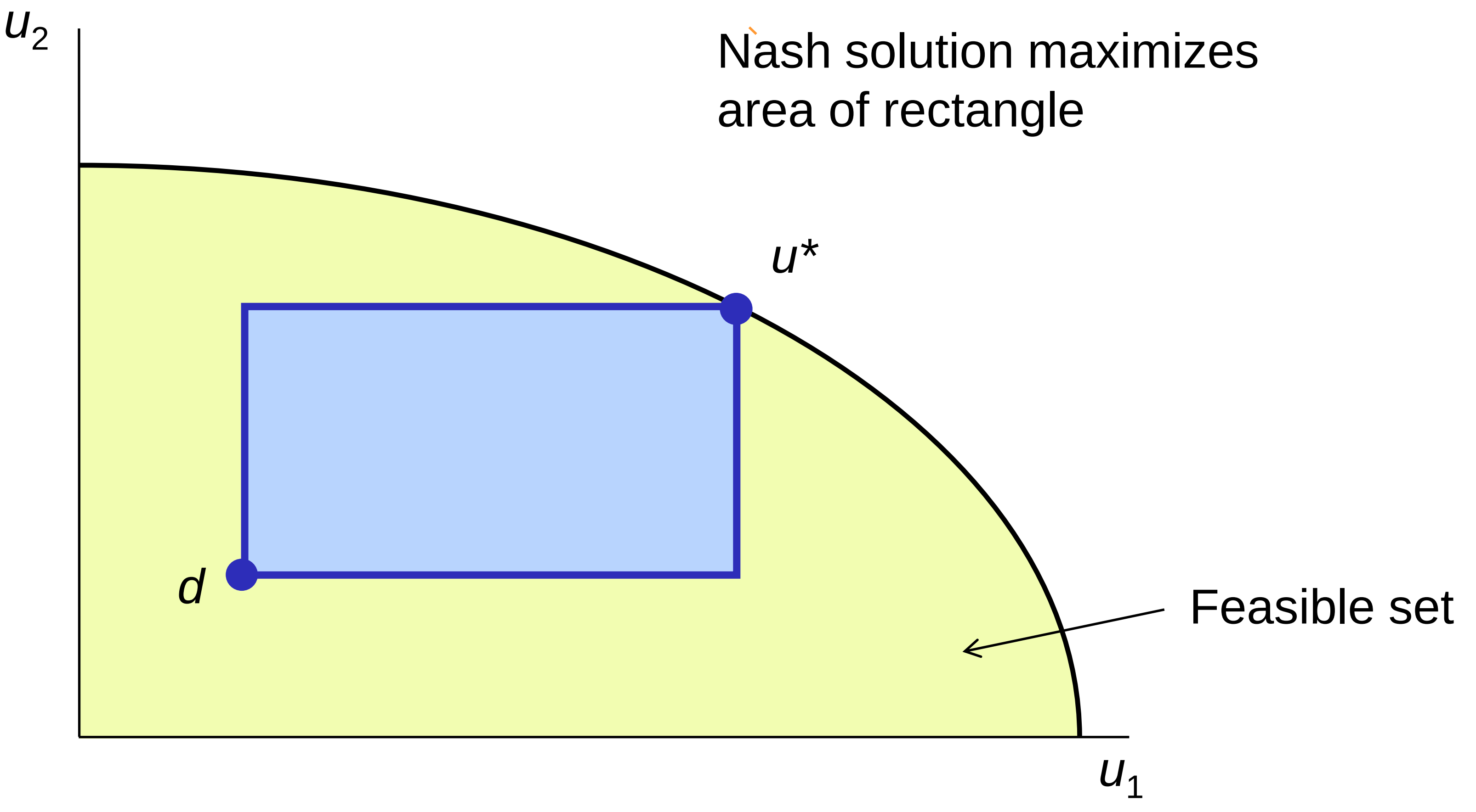


Nash solution maximizes area of rectangle

Proportional Fairness

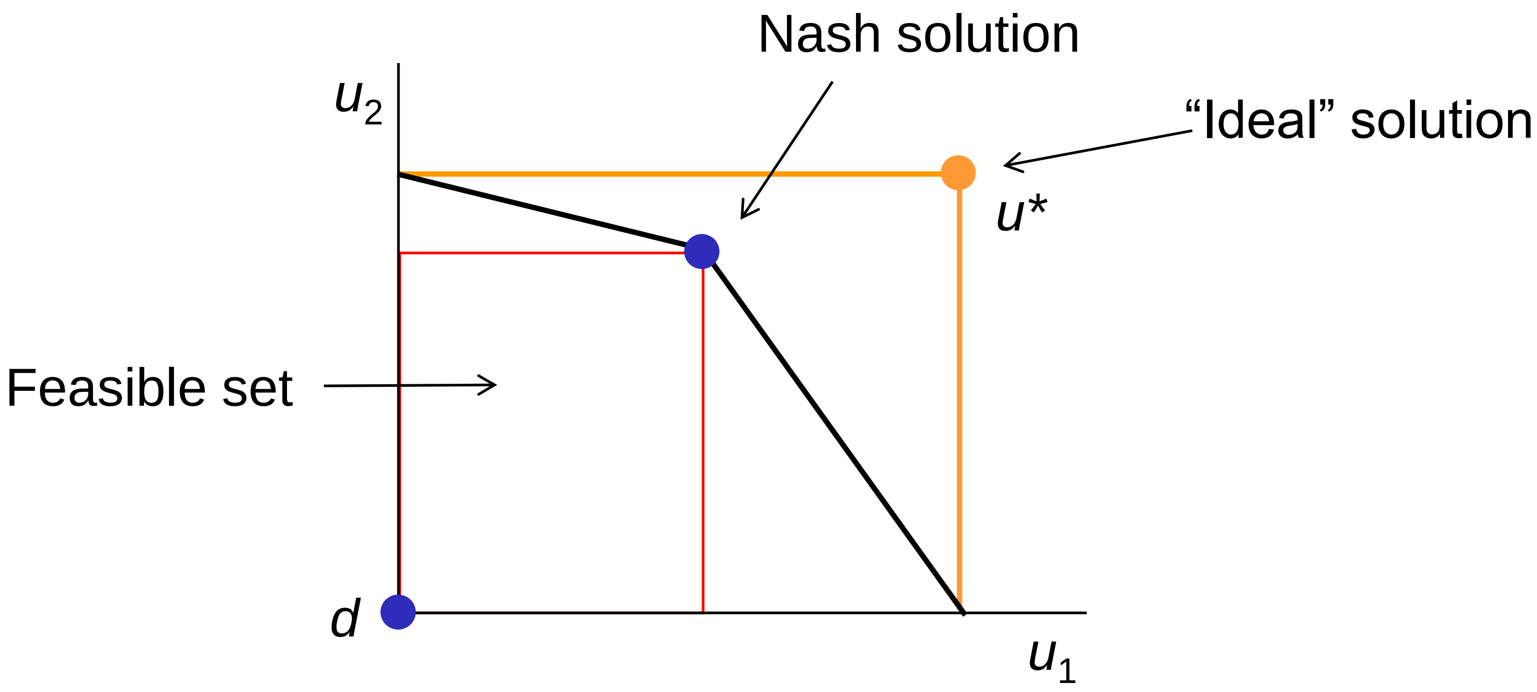


Proportional Fairness



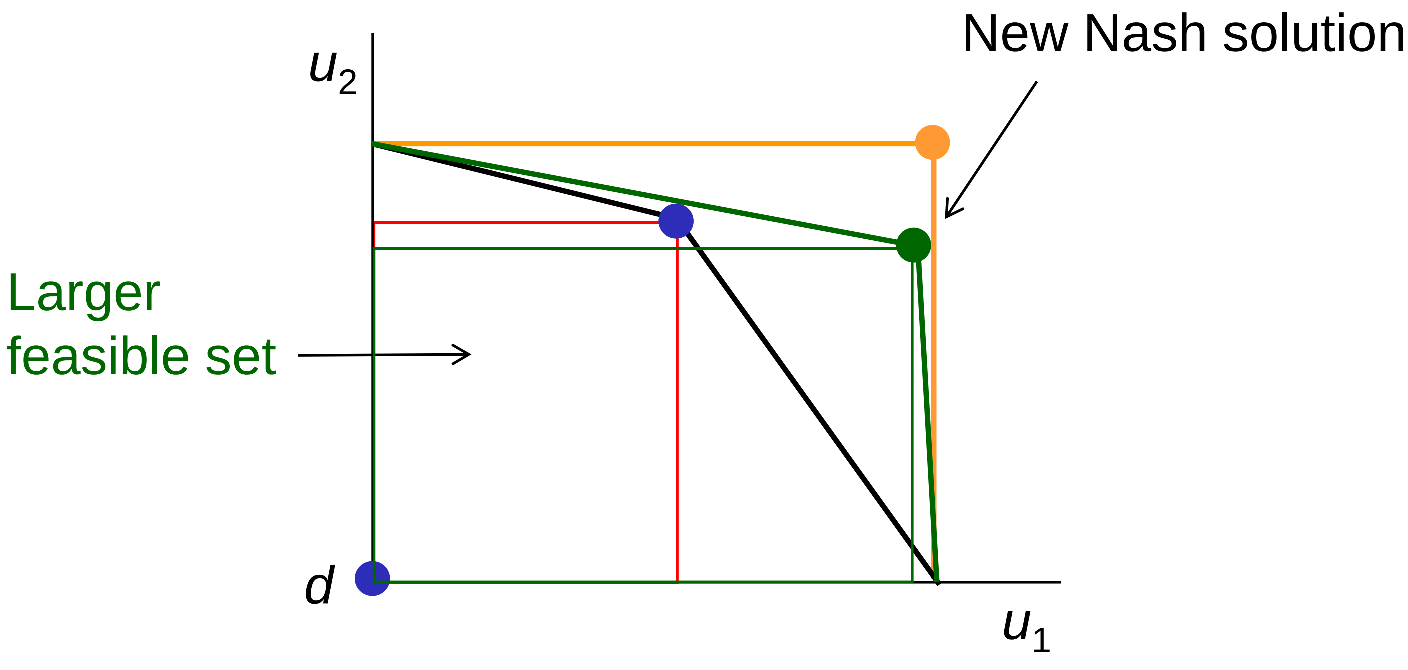
Kalai-Smorodinsky Bargaining

- Begins with a critique of the Nash bargaining solution.



Kalai-Smorodinsky Bargaining

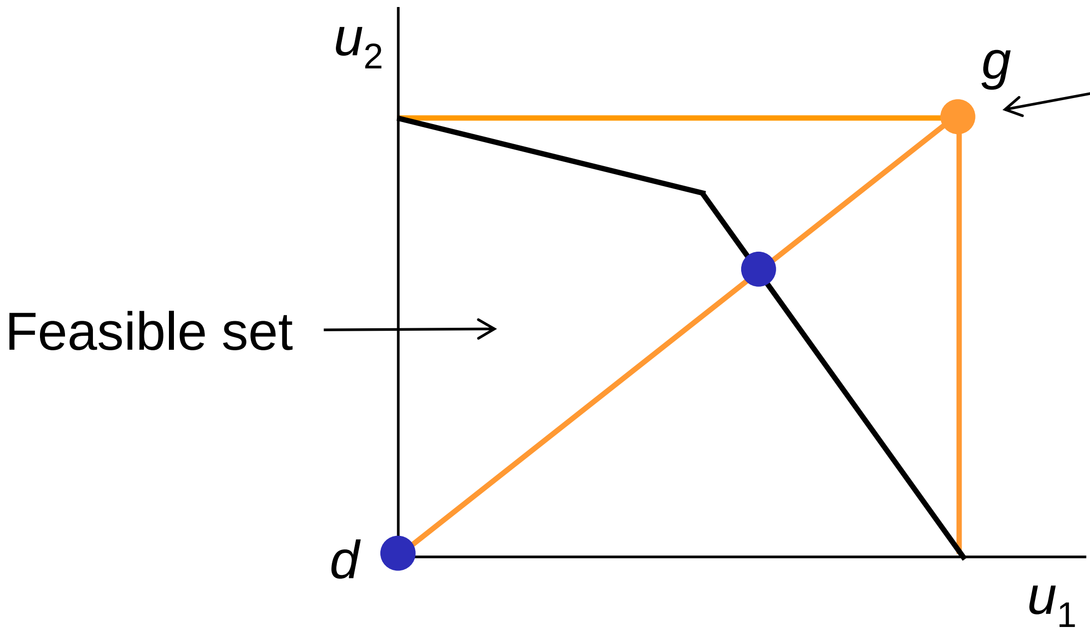
- Begins with a critique of the Nash bargaining solution.
- The new Nash solution is **worse** for player 2 even though the feasible set is **larger**.



Kalai-Smorodinsky Bargaining

- **Proposal**: Bargaining solution is optimal point on the line segment from *d* to ideal solution.
- The players receive an equal fraction of their possible utility gains.

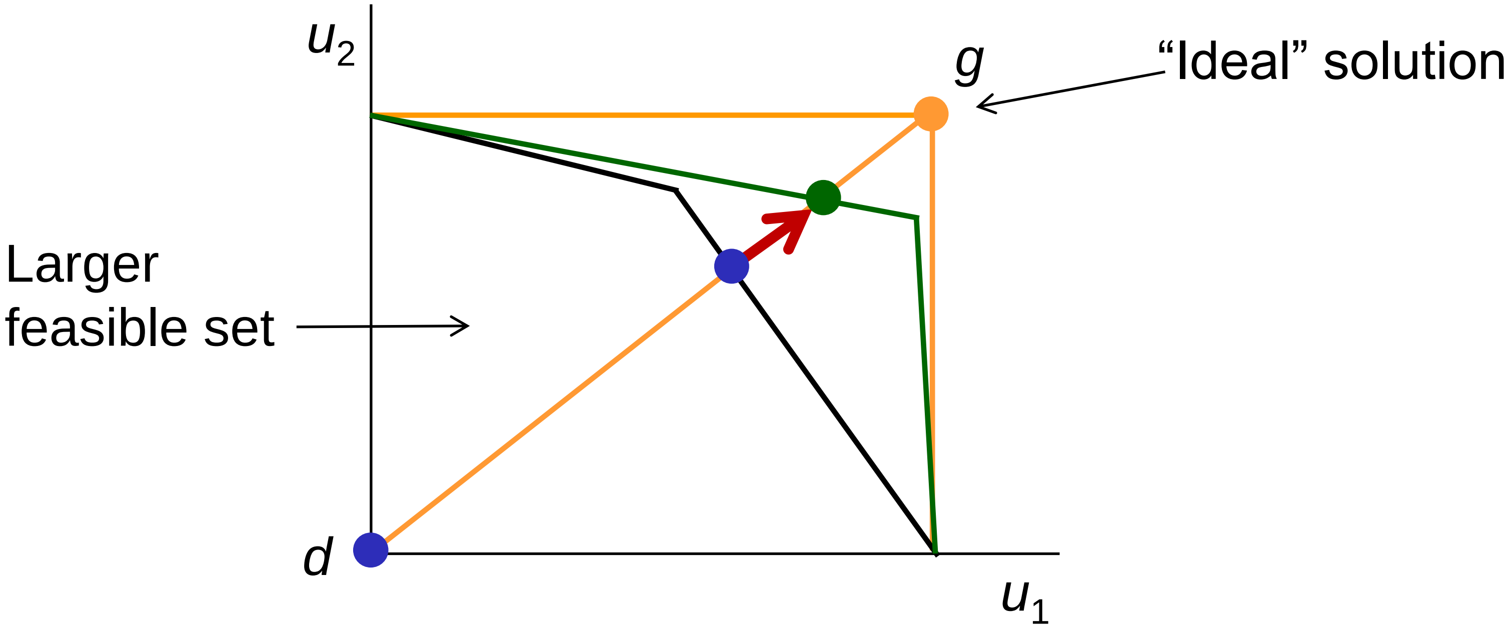
Kalai & Smorodinksy 1975



Kalai-Smorodinsky Bargaining

Rationale

- Satisfies monotonicity, unlike Nash solution.
- Bargaining justification.
- Perhaps suitable for wage, price negotiation.



Kalai-Smorodinsky Bargaining

Possible problems

- May not be ethical to allocate utility in proportion to one's potential.
- For example, when allocating resources to those with minor ailments vs chronic diseases.

Statistical Fairness Metrics

- Widely discussed in AI.
- Intended to measure bias across groups.
- Most are based on statistical measures of classification error.
- Utility vector $\boldsymbol{u}$ is now vector $\boldsymbol{\delta}$ of yes-no decisions.
- For example: mortgage loans, job interviews, parole.

Rationale

- Unjustified bias against certain groups generally seen as inherently unfair.
- Bias may also incur legal problems.

Statistical Fairness Metrics

Notation

- TP = number of true positives (correct yes's)
- FP = number of false positives (incorrect yes's).
- TN = number of true negatives (correct no's).
- FN = number of false negatives (incorrect no's).

Basic model

- Maximize **accuracy**, perhaps $$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$$

  …subject to a **bound** on a bias SWF.
- Bias measured by **comparing various statistics** across 2 groups (a protected group and everyone else).

Statistical Fairness Metrics

Demographic parity

- Compare $\dfrac{\mathrm{TP}+\mathrm{FP}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$ across 2 groups

$$W(\boldsymbol{\delta}) = 1 - |B(\boldsymbol{\delta})|, \text{ where } B(\boldsymbol{\delta}) = \underbrace{\frac{1}{|N|}}_{\text{Majority group}} \sum_{i \in N} \delta_i - \underbrace{\frac{1}{|N'|}}_{\text{Minority group}} \sum_{i \in N'} \delta_i$$

Rationale

- Equality of outcomes.

Possible problem

- Can discriminate against a minority group that is more qualified than majority group.

Dwork et al. 2012

Statistical Fairness Metrics

Equalized odds

Equality of opportunity

- Compare $\dfrac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$ and $\dfrac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}$ across 2 groups

$$B(\boldsymbol{\delta}) = \frac{\sum_{i\in N} a_i\delta_i}{\sum_{i\in N} a_i} - \frac{\sum_{i\in N'} a_i\delta_i}{\sum_{i\in N'} a_i}$$ and $$B(\boldsymbol{\delta}) = \frac{\sum_{i\in N}(1-a_i)\delta_i}{\sum_{i\in N}(1-a_i)} - \frac{\sum_{i\in N'}(1-a_i)\delta_i}{\sum_{i\in N'}(1-a_i)}$$

Rationale

- Compares fraction of **qualified** (or unqualified) persons selected.

Possible problem

- Considers only **yes** (or only **no**) decisions.

Hardt et al. 2016

Statistical Fairness Metrics

Accuracy parity

- Compare $\dfrac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$ across 2 groups.

$$B(\boldsymbol{\delta}) = \frac{1}{|N|}\sum_{i\in N}\big(a_i\delta_i + (1-a_i)(1-\delta_i)\big) - \frac{1}{|N'|}\sum_{i\in N'}\big(a_i\delta_i + (1-a_i)(1-\delta_i)\big)$$

Rationale

Berk et al. 2018

- Compares overall accuracy.
- Only one comparison needed, rather than 2 as in equalized odds.

Possible problem

- Less popular in applications.

Statistical Fairness Metrics

Predictive rate parity

- Compare $\dfrac{\text{TP}}{\text{TP} + \text{FP}}$ across 2 groups.

$$B(\boldsymbol{\delta}) = \frac{\sum_{i \in N} a_i \delta_i}{\sum_{1 \in N} \delta_i} - \frac{\sum_{i \in N'} a_i \delta_i}{\sum_{1 \in N'} \delta_i}$$

Rationale

- Compares what fraction of selected individuals **should** have been selected.

Dieterich et al. 2016

Problem

- Poses very difficult nonconvex discrete optimization problem.
- Unclear what justifies the computational burden.

Statistical Fairness Metrics

Counterfactual fairness

Rationale

- Attempts to determine whether the decision for minority individuals would have been different if they were majority individuals.
- Computes conditional probabilities on Bayesian (causal) networks.

Kusner et al. 2017, Russell et al. 2017

Problems

- Unclear if data are available to allow a reliable determination of causality.
- Unclear how to embed this into a social welfare optimization model.

Statistical Fairness Metrics

General problems

- Yes-no outcomes (δ) provide a **limited perspective** on utility consequences (***u***)***.***
- **No consensus** on which bias metric $B(\delta)$, if any, is suitable for a given context. Bias metrics were developed to measure predictive accuracy, not fairness.
- No principle for **balancing** equity and efficiency.
- No clear principle for **selecting protected groups** (N), unless one simply selects those protected by law.

Threshold Methods with Maximin

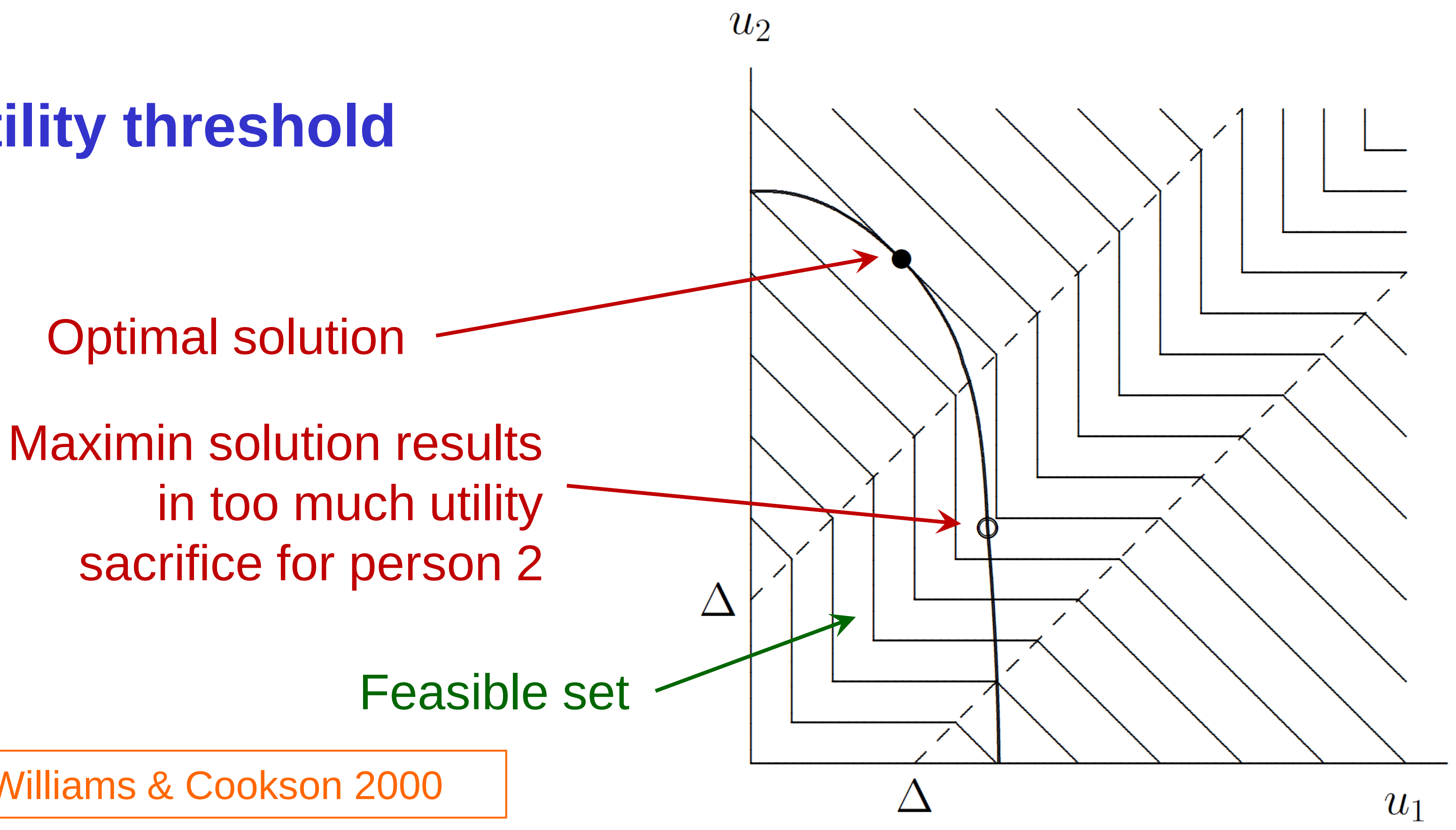
Combining utility and maximin

- **Utility threshold:** Use a maximin criterion until the utility cost becomes too great, then switch a utilitarian criterion.
- **Equity threshold:** Use a utilitarian criterion until the inequity becomes too great, then switch to a maximin criterion.

Williams & Cookson 2000

Threshold Methods with Maximin

Utility threshold



Williams & Cookson 2000

$$W(u_1, u_2) = \begin{cases} u_1 + u_2, & \text{if } |u_1 - u_2| \geq \Delta \\ 2\min\{u_1, u_2\} + \Delta, & \text{otherwise} \end{cases}$$

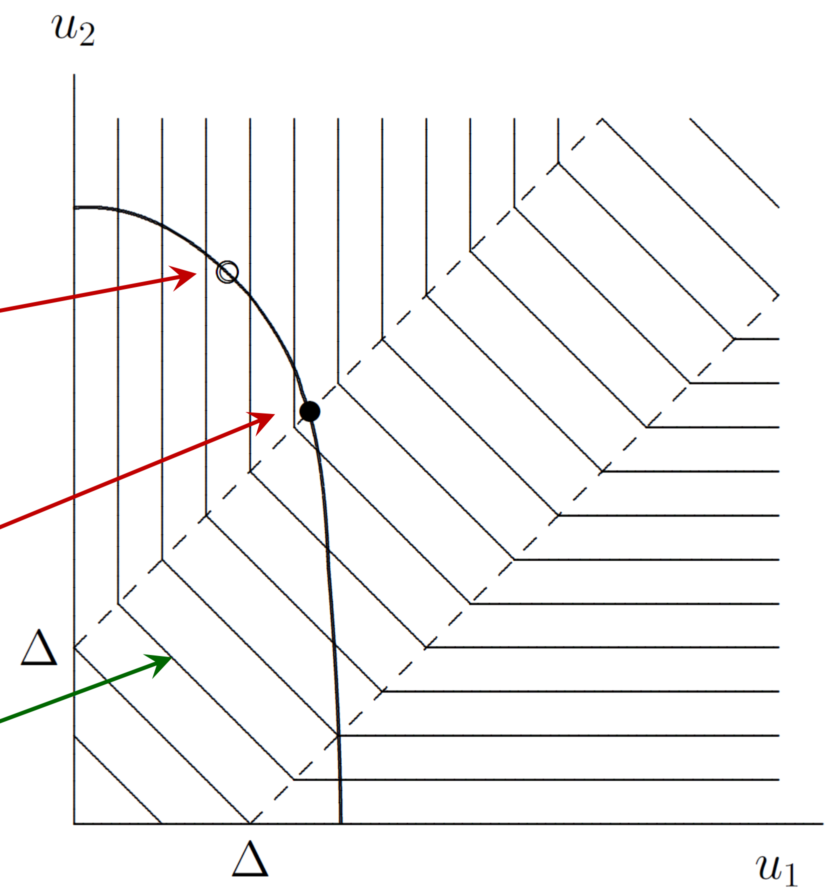
Threshold Methods with Maximin

Equity threshold

Utilitarian solution leaves person 1 overly deprived

Optimal solution

Feasible set



Williams & Cookson 2000

$$W(u_1, u_2) = \begin{cases} 2\min\{u_1, u_2\} + \Delta, & \text{if } |u_1 - u_2| \geq \Delta \\ u_1 + u_2, & \text{otherwise} \end{cases}$$

Threshold Methods with Maximin

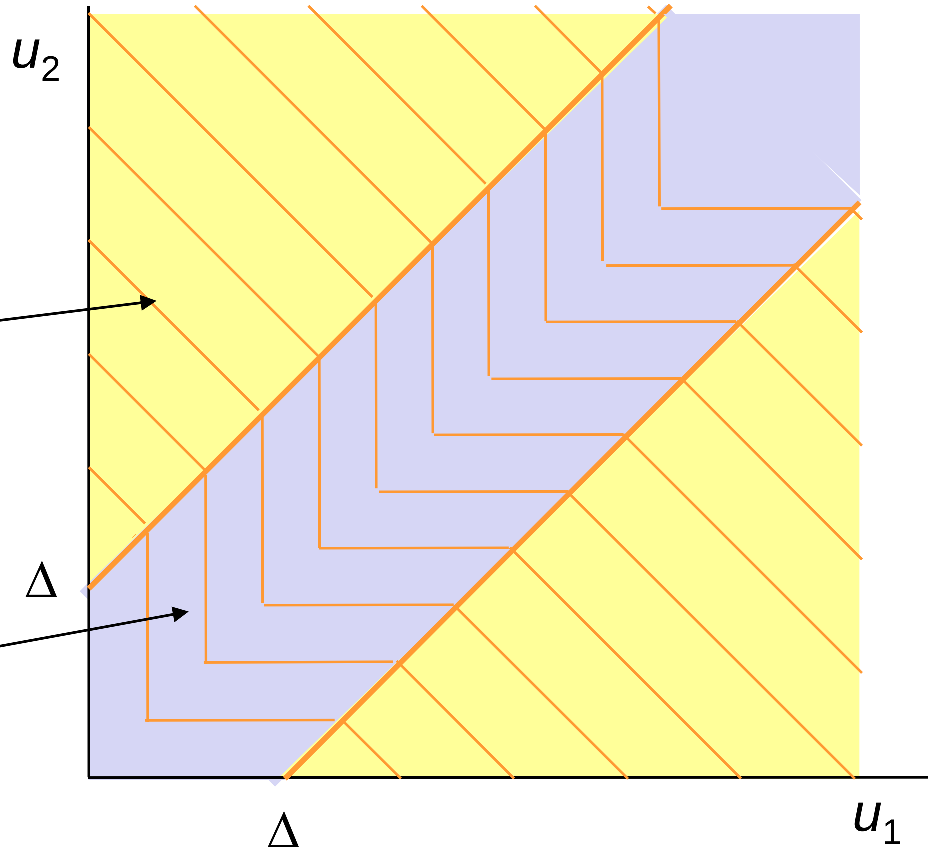
Combining utility and maximin

Rationale

- **Utility threshold:** Suitable when equity is the main consideration, but without excessive utility sacrifice.
  - As in medical applications, politically sensitive contexts.
- **Equity threshold:** Suitable when utility is the main consideration, but without sacrificing basic equity.
  - As in telecommunications, disaster management, traffic control.

Utility Threshold with Maximin

Use maximin criterion until it results in excessive sacrifice by some individuals



Utilitarian region

$$u_1 + u_2$$

Maximin region

$$2\min\{u_1, u_2\} + \Delta$$

Ensures continuous contours

Utility Threshold with Maximin

Generalization to *n* persons

$$W(\boldsymbol{u}) = (n-1)\Delta + \sum_{i=1}^{n} \max\left\{u_i - \Delta, u_{\min}\right\}$$

$$\text{where } u_{\min} = \min_i \{u_i\}$$

JH & Williams 2012

Rationale

- Utilities within Δ of the lowest are in the **fair region**.
- Utilities in fair region are **equated** with smallest utility.
  - In effect, this gives weight to lowest utility equal to number of utilities in the fair region.

i-th largest utility

$$W(\boldsymbol{u}) = t(\boldsymbol{u})u_{\langle 1\rangle} + \big(t(\boldsymbol{u})-1\big)\Delta + \sum_{i=t(\boldsymbol{u})+1}^{n} u_{\langle i\rangle}$$

Number of utilities in fair region

Utility Threshold with Maximin

Generalization to *n* persons

$$W(\boldsymbol{u}) = (n-1)\Delta + \sum_{i=1}^{n} \max\left\{u_i - \Delta, u_{\min}\right\}$$

where $u_{\min} = \min_i \{u_i\}$

JH & Williams 2012

Rationale

- Trade-off parameter Δ has a **practical interpretation**.
- Δ is chosen so that individuals in fair region are sufficiently deprived to **deserve priority**.
- $\Delta = 0$ corresponds to utilitarian criterion, $\Delta = \infty$ to maximin.

Utility Threshold with Maximin

Utility threshold

Model

$$\max_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{\delta},\boldsymbol{v},w,z} \left\{ n\Delta + \sum_i v_i \;\middle|\; \begin{array}{c} u_i - \Delta \le v_i \le u_i - \Delta\delta_i, \text{ all } i \\ w \le v_i \le w + (M-\Delta)\delta_i, \text{ all } i \\ u_i - u_i \le M, \text{ all } i,j \\ u_i \ge 0, \ \delta_i \in \{0,1\}, \text{all } i \\ (\boldsymbol{u},\boldsymbol{x}) \in S \end{array} \right\}$$

- Tractable MILP model.
- Model is **sharp** without ($\boldsymbol{u}$, $\boldsymbol{x}$) $\in S$.
- Easily generalized to differently-sized **groups** of individuals.

JH & Williams 2012

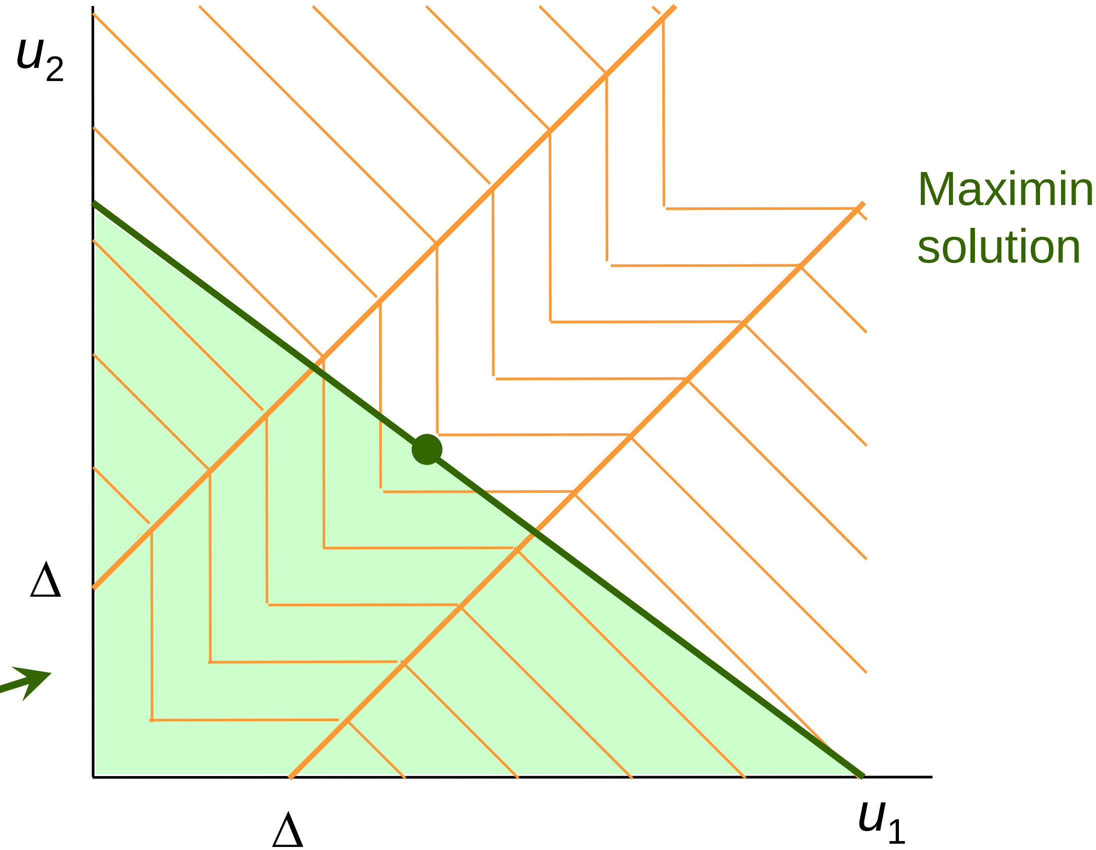
Utility Threshold with Maximin

Theorem. When maximizing the SWF subject to a **budget constraint**, the optimal solution is purely **maximin** or **purely utilitarian**.

Purely maximin if

$$\Delta \geq B\left(\frac{1}{a_{\langle 1 \rangle}} - \frac{n}{\sum_i a_i}\right)$$

Here, patients have **similar** treatment costs, or Δ is **large**.



Elçi, JH, and Zhang 2021

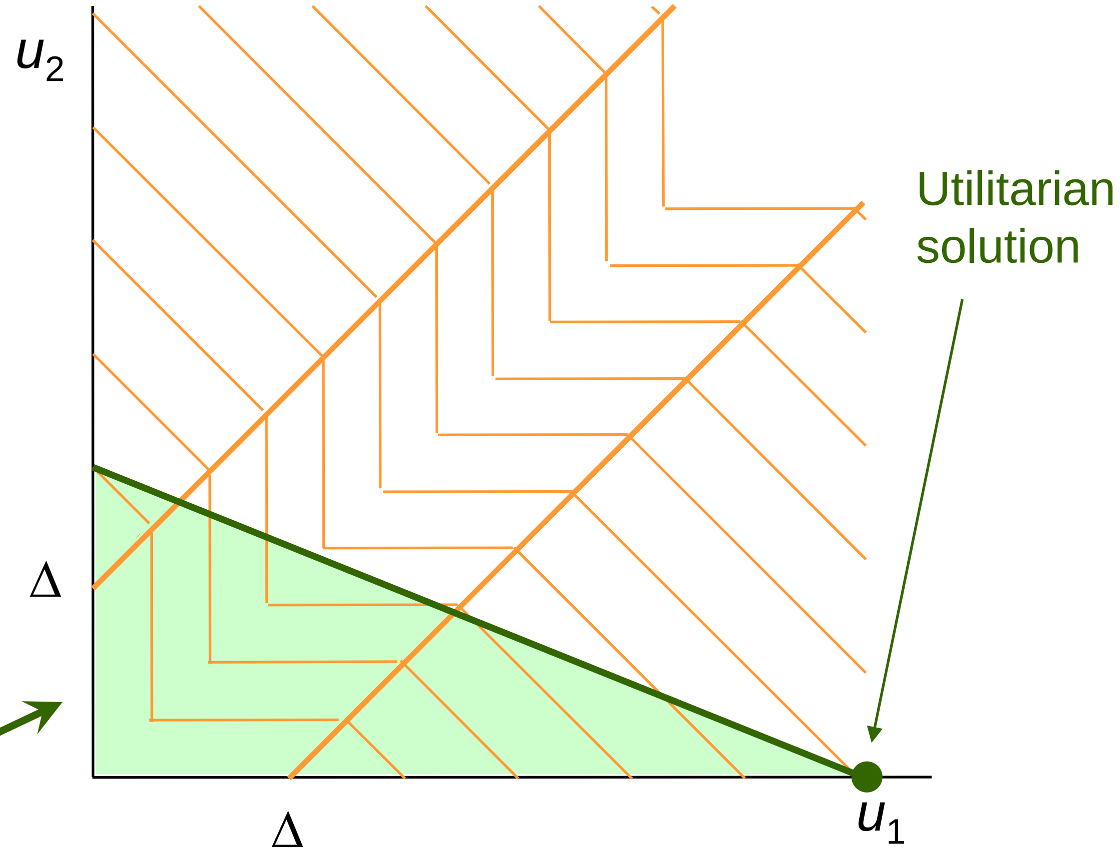
Utility Threshold with Maximin

Theorem. When maximizing the SWF subject to a **budget constraint**, the optimal solution is purely **maximin** or **purely utilitarian**.

Purely utilitarian if

$$\Delta \le B\left(\frac{1}{a_{\langle 1 \rangle}} - \frac{n}{\sum_i a_i}\right)$$

Here, patients have **very different** treatment costs, or Δ is **small**.

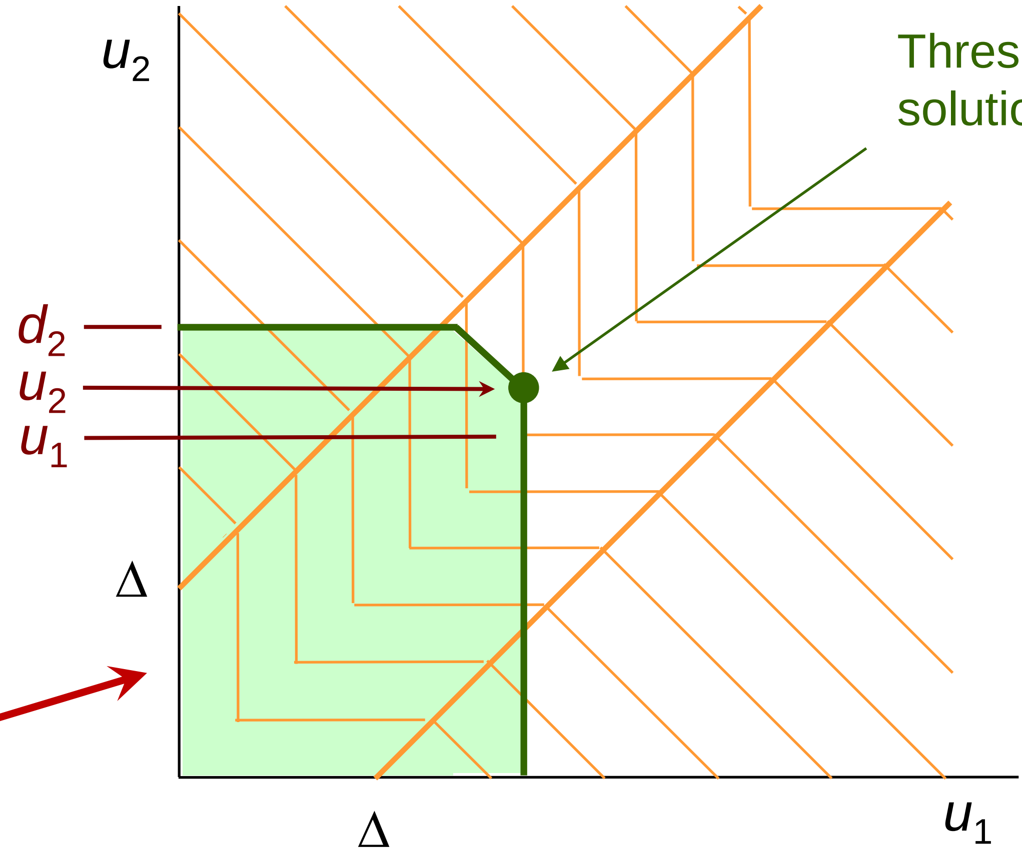


Elçi, JH, and Zhang 2021

Utility Threshold with Maximin

Theorem. When maximizing the SWF subject to a **budget constraint and upper bounds** d_i at most one utility is **strictly between** its upper bound and the smallest utility.

Here, **one** utility u_2 is **strictly between** upper bound d_2 and the smallest utility u_1.



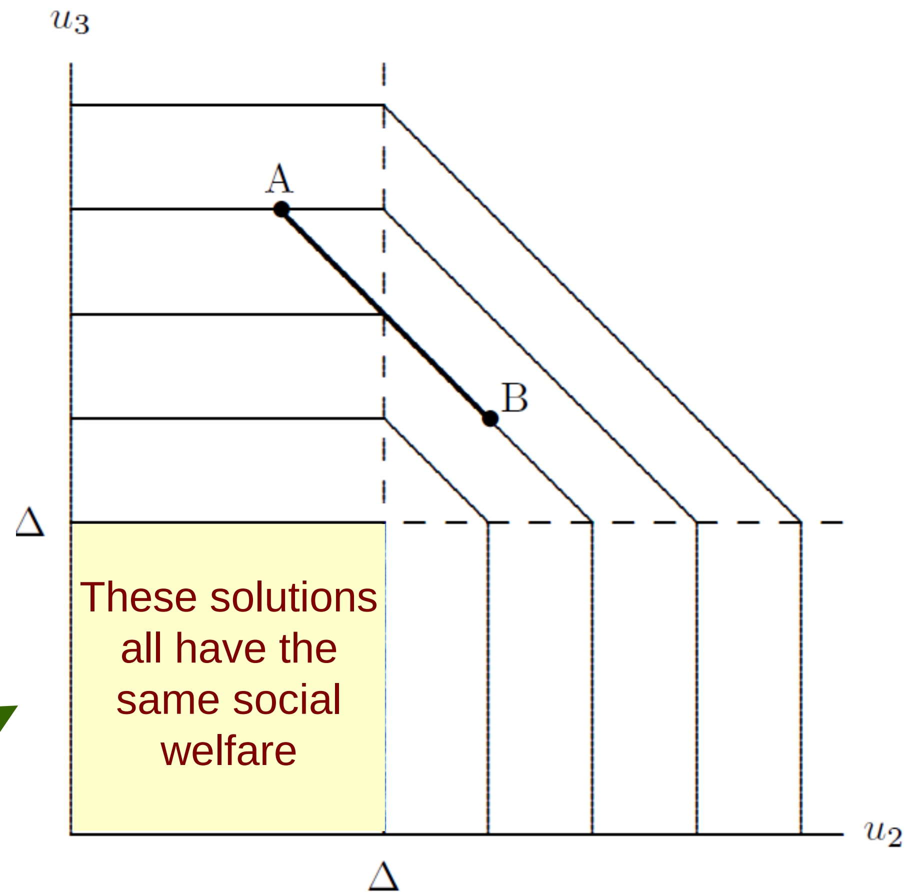
Elçi, JH, and Zhang 2021

Utility Threshold with Maximin

Problem

Too many solutions with different equity properties have the same social welfare.

...because we combine utility with **maximin**



3-person example.
Contours of $W(0,u_1,u_2)$

Equity Threshold with Maximin

Generalization to *n* persons

$$W(\boldsymbol{u}) = n\Delta + \sum_{i=1}^{n} \min\{u_i - \Delta, u_{min}\}$$

Chen & JH 2021

Rationale

- Utilities more than Δ above the lowest are in the **fair region**.
- Δ is chosen so that well-off individuals (those in fair region) **do not deserve more utility** unless smaller utilities are also increased.
- Suitable when **efficiency** is the initial concern, but one does not want to create **excessive inequality** (traffic management, telecom, disaster recovery).

Equity Threshold with Maximin

Equity threshold

Model

$$\max_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},w,z} \left\{ n\Delta + \sum_i v_i \;\middle|\; \begin{array}{c} v_i \le w \le u_i, \text{ all } i \\ v_i \le u_i - \Delta, \text{ all } i \\ w \ge 0, v_i \ge 0, \text{ all } i \\ (\boldsymbol{u}, \boldsymbol{x}) \in S \end{array} \right\}$$

Chen & JH 2021

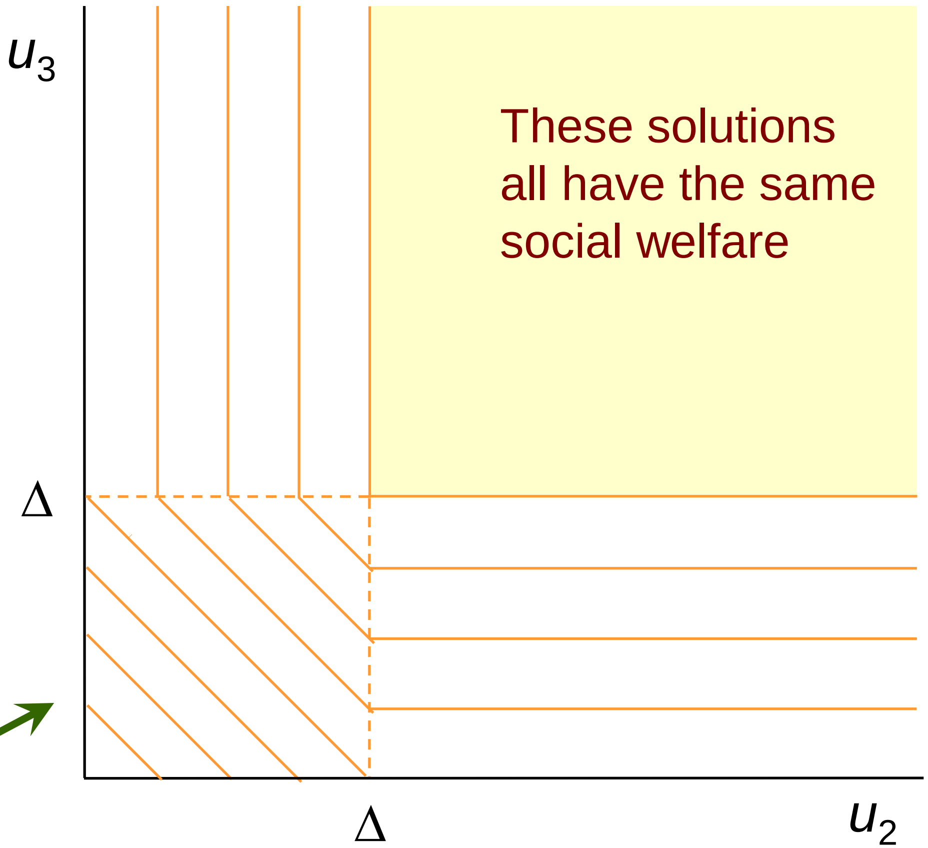
- Linear model.
- Easily generalized to differently-sized **groups** of individuals.

Equity Threshold with Maximin

Problem

Too many solutions with different equity properties have the same social welfare.

...because we combine utility with **maximin**



3-person example.
Contours of $W(0,u_1,u_2)$

Utility Theshold with Leximax

Maximize sequence of SWFs

- Each SWF $F_1(\boldsymbol{u})$, … $F_n(\boldsymbol{u})$ combines equity and utility.
  - Max $F_1(\boldsymbol{u})$ to obtain $u^*_{\langle 1\rangle}$.
  - Then max $F_2(\boldsymbol{u})$ with $u_{\langle 1\rangle} = u^*_{\langle 1\rangle}$ to obtain $u^*_{\langle 2\rangle}$, etc.
  - $(\boldsymbol{u}^*_{\langle 1\rangle}, \ldots, \boldsymbol{u}^*_{\langle n\rangle})$ is socially optimal solution.

Rationale

- Sensitive to equity concerns of disadvantaged parties other than the very worst off.

Chen & JH 2021

Utility Theshold with Leximax

The SWFs

- $W_1(\boldsymbol{u})$ is the utility threshold SWF defined earlier.
- W_k for $k \geq 2$ is

$$W_k(\boldsymbol{u}) = \sum_{i=1}^{k-1}(n-i+1)u_{\langle i\rangle} + (n-k+1)\min\left\{u_{\langle 1\rangle} + \Delta, u_{\langle k\rangle}\right\} + \sum_{i=k}^{n}\max\left\{0,\ u_{\langle i\rangle} - u_{\langle 1\rangle} - \Delta\right\}$$

  where $u_{\langle 1\rangle}, \ldots, u_{\langle n\rangle}$ are $u_1, \ldots, u_n$ in nondecreasing order.

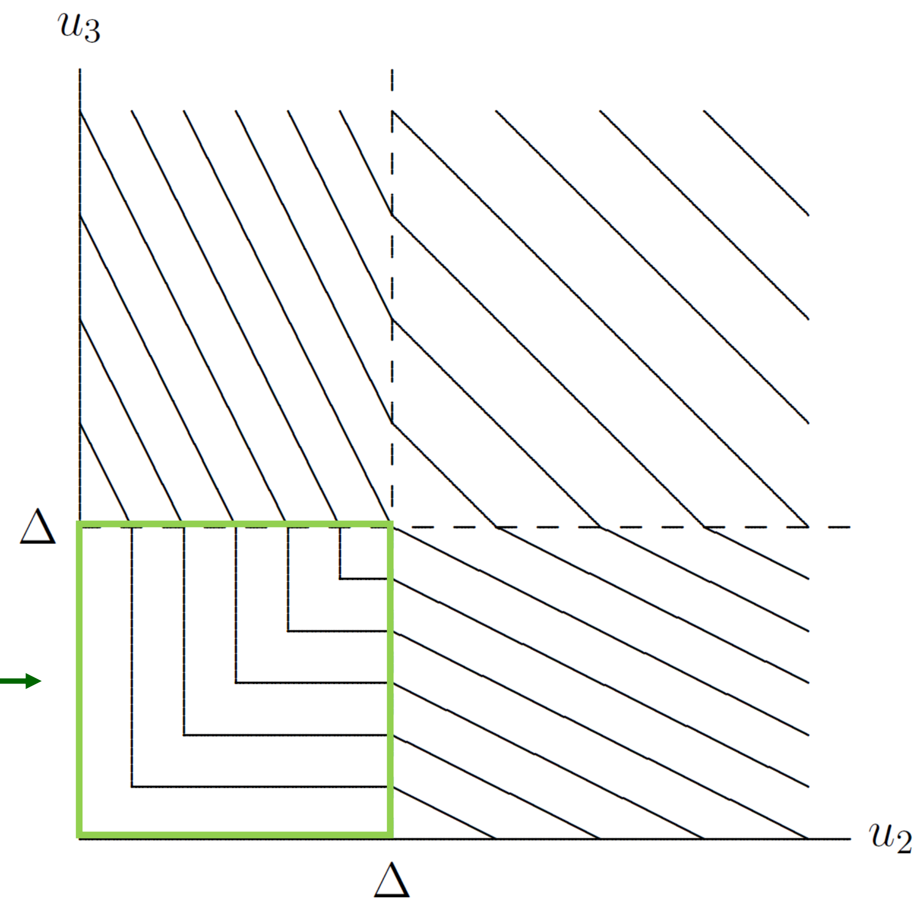
- In $F_k(\boldsymbol{u})$, $u_{\langle k\rangle}$ receives weight $n-k+1$, and $u_{\langle i\rangle}$ for $i > k$ weight 1.
- So, less disadvantaged parties receive less weight.
- When $k \geq 2$, weights cannot depend on $\boldsymbol{u}$ (e.g., on number of utilities in fair region), else the SWF is discontinuous.

Utility Theshold with Leximax

Example: 3 persons

Contours for $F_2(0,u_2,u_3)$

Model is sensitive to equity of all persons in maximin region



Utility Theshold with Leximax

Theorem. In a socially optimal solution subject to a **budget constraint**, solution may be **neither utilitarian nor maximin**.

Theorem. In a socially optimal solution subject to a budget constraint and **bounds**, **several** utilities may lie strictly between their upper bounds and the smallest utility.

Elçi, JH, and Zhang 2021

Utility Theshold with Leximax

Model (MILP for W_k)

$$\max_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{\delta},\boldsymbol{\epsilon}\\ \boldsymbol{v},w,\sigma,z}} \left\{ z \;\middle|\; \begin{array}{l} z \leq (n-k+1)\sigma + \sum_{i\in I_k} v_i \\ 0 \leq v_i \leq M\delta_i, \;\; i \in I_k \\ v_i \leq u_i - \bar{u}_{i_1} - \Delta + M(1-\delta_i), \;\; i \in I_k \\ \sigma \leq \bar{u}_{i_1} + \Delta \\ \sigma \leq w \\ w \leq u_i, \; i \in I_k \\ u_i \leq w + M(1-\epsilon_i), \; i \in I_k \\ \sum_{i\in I_k} \epsilon_i = 1 \\ w \geq \bar{u}_{i_{k-1}} \\ u_i - \bar{u}_{i_1} \leq M, \; i \in I_k \\ \delta_i, \epsilon_i \in \{0,1\}, \; i \in I_k \end{array} \right\}$$

where $\bar{u}_{i_k}$ is the value of the smallest utility in the optimal solution of the kth MILP model, and $I = \{1,\ldots,n\} \setminus \{i_1,\ldots,i_{k-1}\}$. The socially optimal solution is $(\bar{u}_1,\ldots,\bar{u}_n)$.

Utility Theshold with Leximax

Theorem. The following inequalities are valid in the MILP model for $F_k(\boldsymbol{u})$.

$$z_k \leq \sum_{i \in I_k} u_i$$

$$z_k \leq (n-k+1)u_i + \beta \sum_{j \in I_k \setminus \{i\}} (u_i - \bar{u}_{i_{k-1}}), \quad i \in I_k$$

where $$\beta = \frac{M - \Delta}{M - (\bar{u}_{i_{k-1}} - \bar{u}_{i_1})}$$

Theshold Methods – Healthcare Example

- Based on budget decisions in UK National Health Service
- Allocate limited treatment resources to disease/prognosis categories of patients.
- Based on cost, number of patients, and QALY estimates with and without treatment.*
- We will compare **2 utility-threshold SWFs**: utility + maximin and sequential utility + leximax.
- Solution time = fraction of second for each value of Δ.

Problem due to JH & Williams 2012

*QALY = quality adjusted life-year. Data reflect a particular situation and are not valid in general. Solutions presented here should not be taken as a general recommendation for healthcare resource allocation, but only as an illustration of social welfare functions.

QALY & cost data

Part 1

Intervention	Cost per person c_i (£)	QALYs gained q_i	Cost per QALY (£)	QALYs without intervention α_i	Subgroup size n_i
Pacemaker for atrioventricular heart block					
Subgroup A	3500	3	1167	13	35
Subgroup B	3500	5	700	10	45
Subgroup C	3500	10	350	5	35
Hip replacement					
Subgroup A	3000	2	1500	3	45
Subgroup B	3000	4	750	4	45
Subgroup C	3000	8	375	5	45
Valve replacement for aortic stenosis					
Subgroup A	4500	3	1500	2.5	20
Subgroup B	4500	5	900	3	20
Subgroup C	4500	10	450	3.5	20
CABG[1] for left main disease					
Mild angina	3000	1.25	2400	4.75	50
Moderate angina	3000	2.25	1333	3.75	55
Severe angina	3000	2.75	1091	3.25	60
CABG for triple vessel disease					
Mild angina	3000	0.5	6000	5.5	50
Moderate angina	3000	1.25	2400	4.75	55
Severe angina	3000	2.25	1333	3.75	60
CABG for double vessel disease					
Mild angina	3000	0.25	12,000	5.75	60
Moderate angina	3000	0.75	4000	5.25	65
Severe angina	3000	1.25	2400	4.75	70

QALY & cost data

Part 2

Intervention	Cost per person c_i (£)	QALYs gained q_i	Cost per QALY (£)	QALYs without intervention α_i	Subgroup size n_i
Heart transplant					
	22,500	4.5	5000	1.1	2
Kidney transplant					
Subgroup A	15,000	4	3750	1	8
Subgroup B	15,000	6	2500	1	8
Kidney dialysis					
Less than 1 year survival					
Subgroup A	5000	0.1	50,000	0.3	8
1-2 years survival					
Subgroup B	12,000	0.4	30,000	0.6	6
2-5 years survival					
Subgroup C	20,000	1.2	16,667	0.5	4
Subgroup D	28,000	1.7	16,471	0.7	4
Subgroup E	36,000	2.3	15,652	0.8	4
5-10 years survival					
Subgroup F	46,000	3.3	13,939	0.6	3
Subgroup G	56,000	3.9	14,359	0.8	2
Subgroup H	66,000	4.7	14,043	0.9	2
Subgroup I	77,000	5.4	14,259	1.1	2
At least 10 years survival					
Subgroup J	88,000	6.5	13,538	0.9	2
Subgroup K	100,000	7.4	13,514	1.0	1
Subgroup L	111,000	8.2	13,537	1.2	1

Theshold Methods – Healthcare Example

Budget constraint

$$\sum_j n_j c_j y_j \leq B$$

Size of treatment group *j*

Unit cost of treatment *j*

Fraction of group treated

Utility function

$$u_i = q_i y_i + \alpha_i$$

Treatment benefit (QALYs)

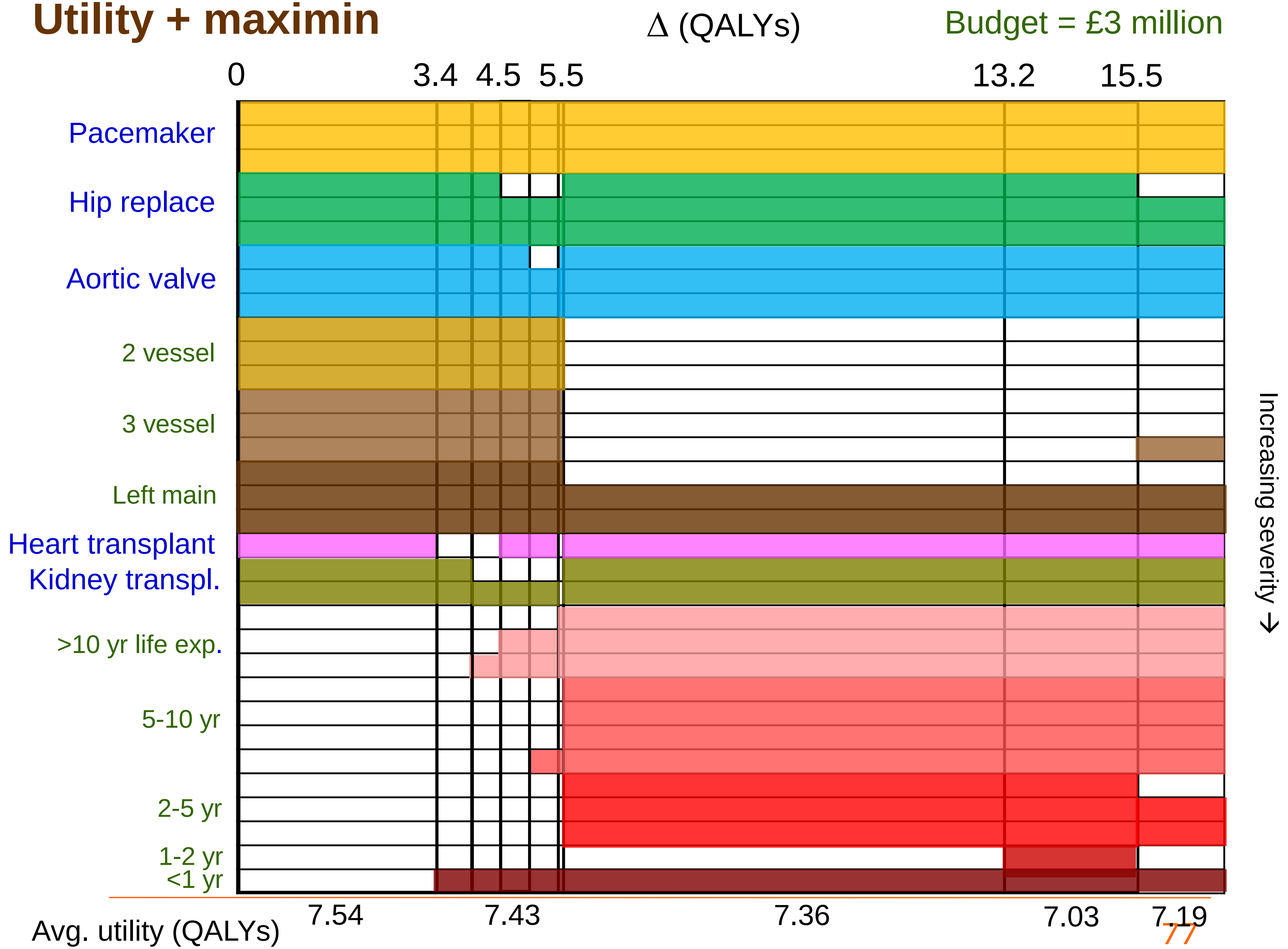
QALYs without treatment

which implies $y_i = (u_i - \alpha_i)/q_i$

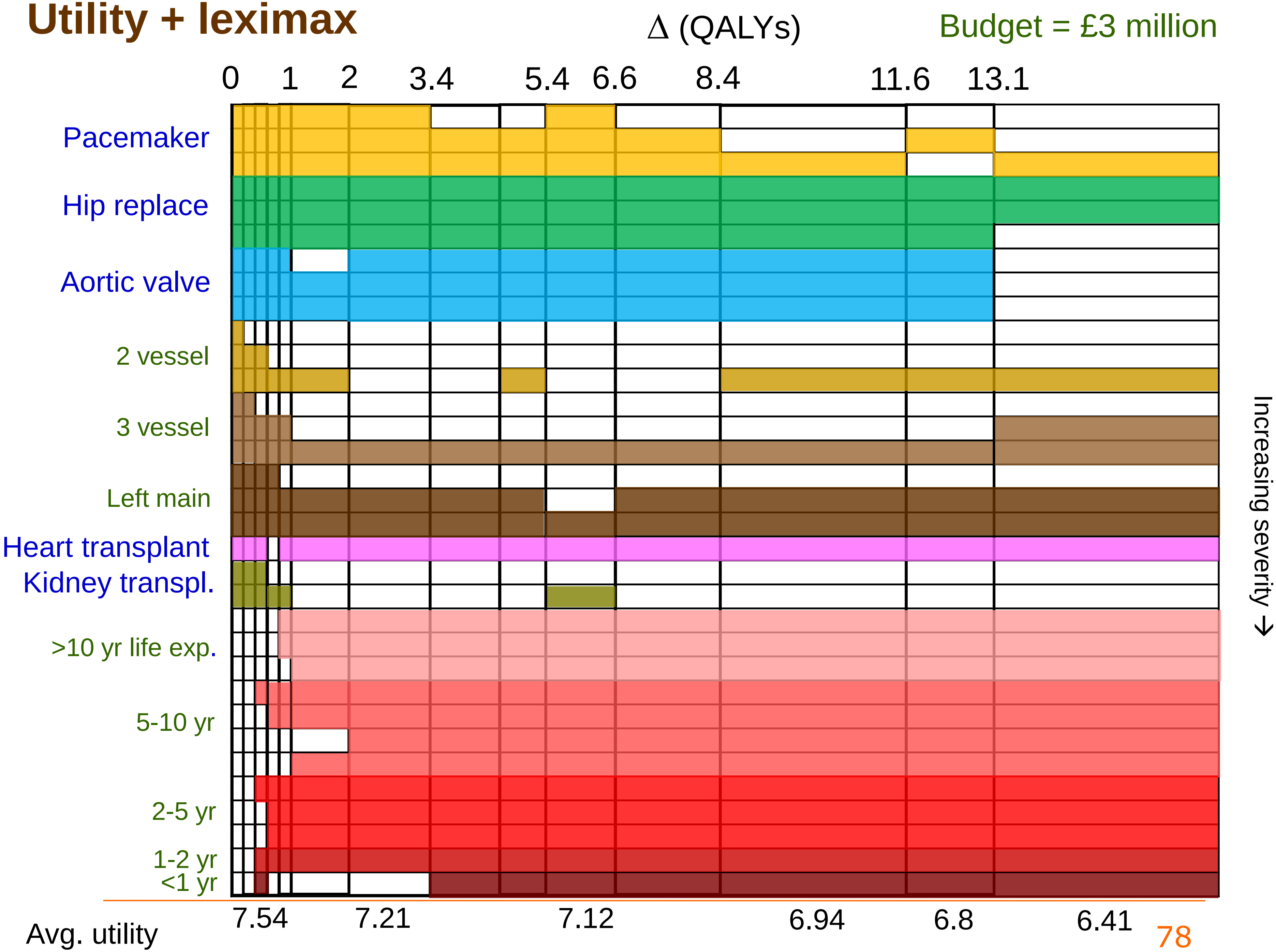
So the optimization problem becomes

$$\max_{\boldsymbol{u}} \left\{ W(\boldsymbol{u}) \;\middle|\; \sum_j \frac{n_j c_j}{q_j} u_j \leq B + \sum_j \frac{n_j c_j \alpha_j}{q_j};\;\; \boldsymbol{\alpha} \leq \boldsymbol{u} \leq \boldsymbol{q} + \boldsymbol{\alpha} \right\}$$

Utility + maximin



Utility + leximax



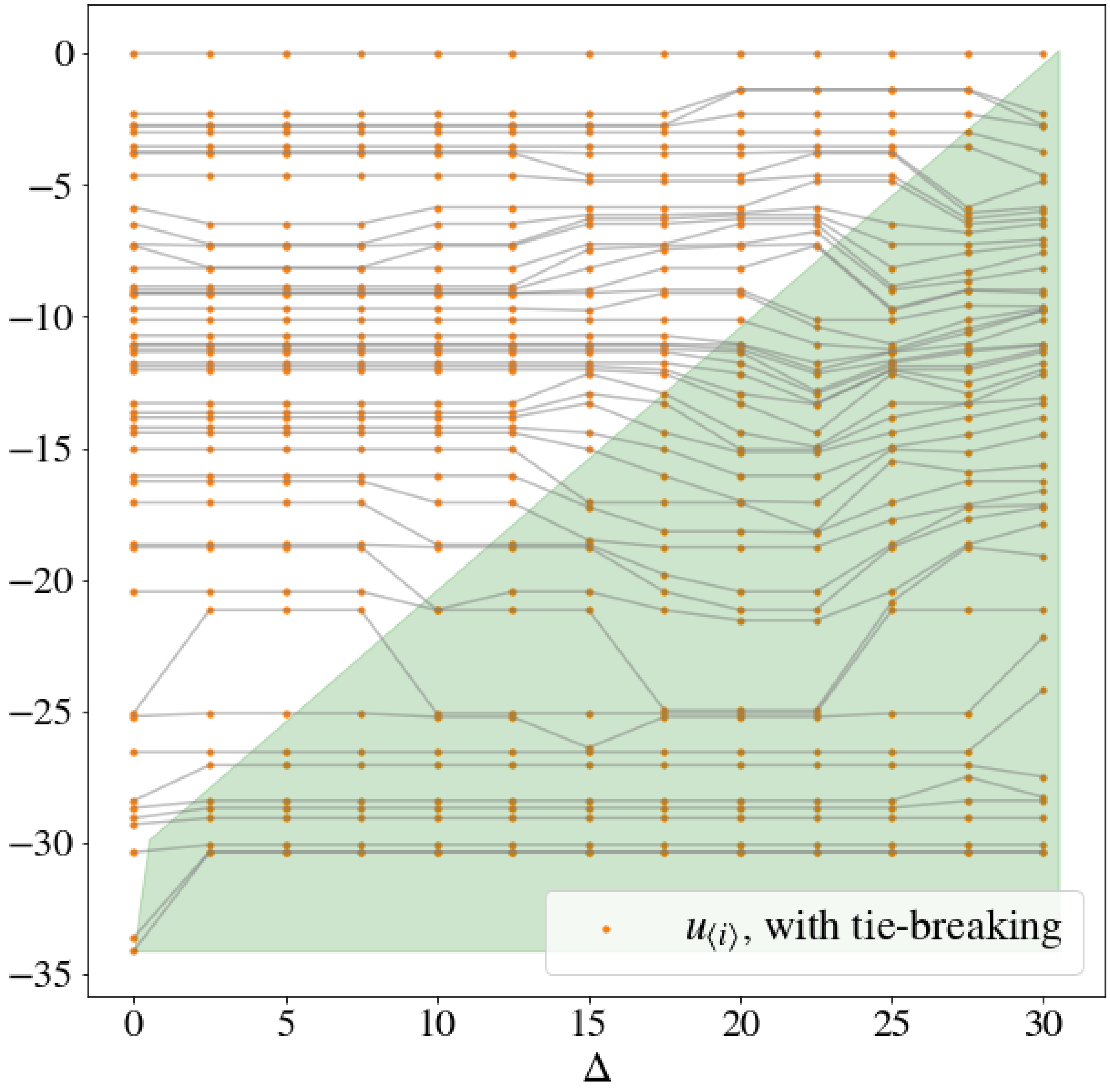
Threshold Methods –
Disaster Preparedness Example

- Select earthquake shelter locations.
- Utility = negative distance of each neighborhood to nearest shelter, subject to limited budget.
- We will compare **2 utility-threshold SWFs**: utility + maximin and sequential utility + leximax.
- 50 neighborhoods, 50 potential shelter locations.
- Solution time = 1 to 18 seconds for each value of Δ.

Problem due to Mostajabdaveh, Gutjahr & Salman 2019

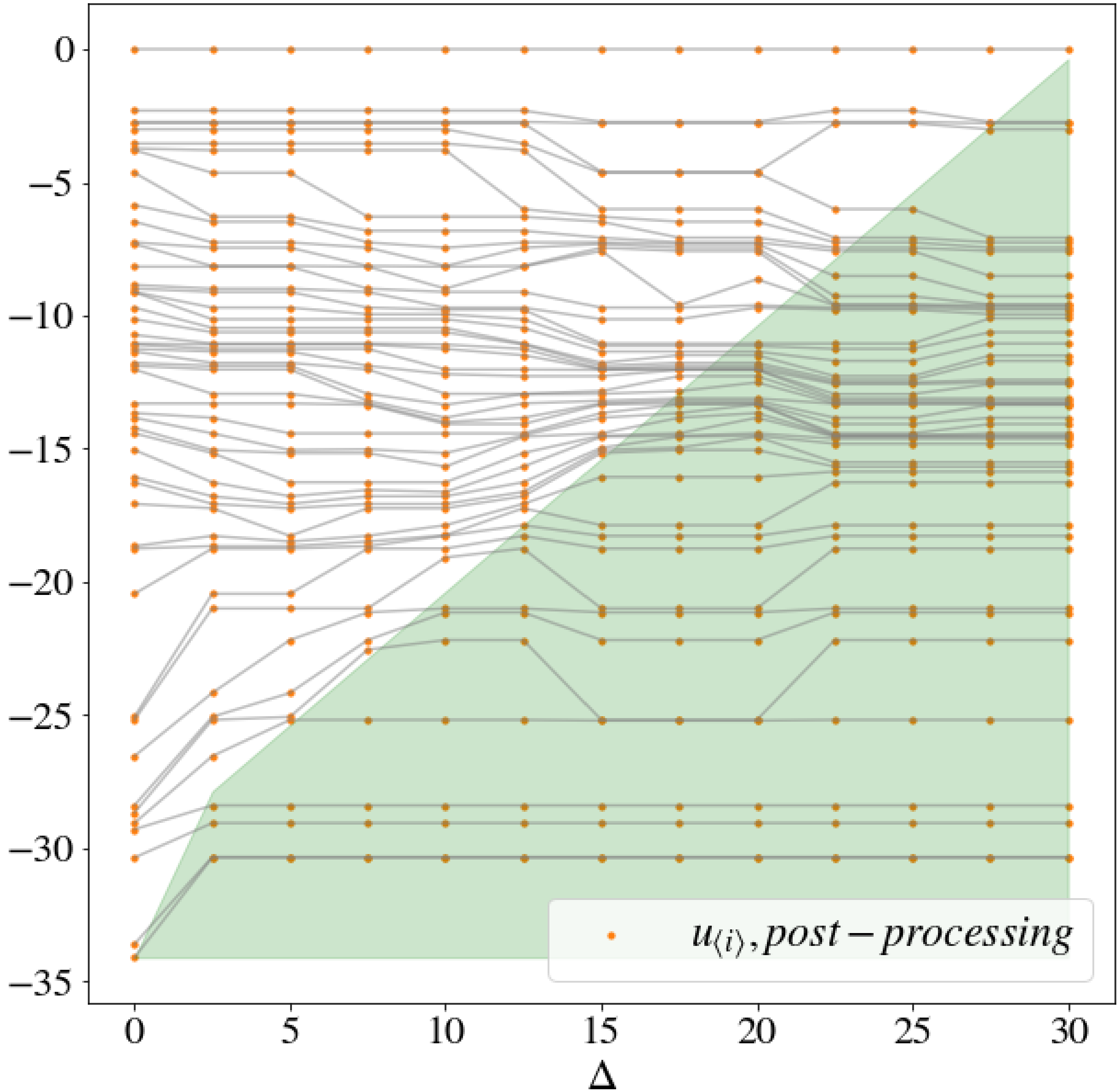
Threshold SWF

Utility + maximin



Threshold SWF

Utility + leximax



References

- J. N. Hooker, Tutorial in equity modeling, CP 2021, video (90 minutes), slides
- V. Chen & J. N. Hooker, Combining leximax fairness and efficiency in a mathematical programming model, *European Journal of Operational Research*, to appear.
- V. Chen & J. N. Hooker, A guide to formulating equity and fairness in an optimization model, submitted, 2021.
- Ö. Elçi, J. N. Hooker, and P. Zhang, Structural characteristics of equitable and efficient distributions, 2021.