Robust dissimilarity comparisons with ordinal outcomes

Francesco Andreoli, University of Verona and LISER
Claudio Zoli, University of Verona

EEA-ESEM 2022, Milan

Part of the research projects *ORDINEQ* (ANR), *IMCHILD* (NORFACE ERA-NET), *PREOPP* (RiBa University of Verona)."

Motivation

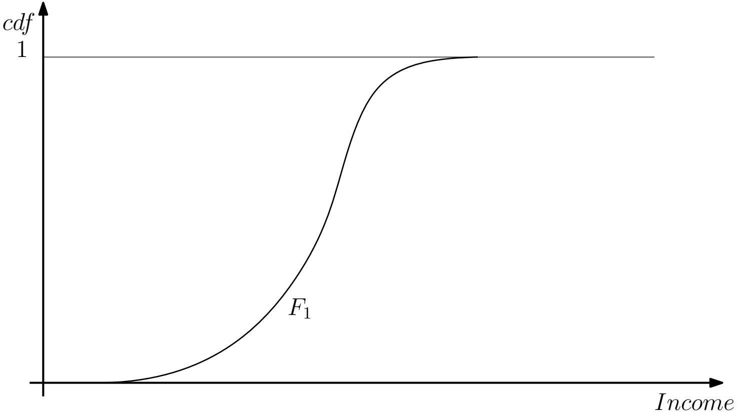
- We analyze the inequalities between distributions (*groups*) of an ordered attribute.
- *Dissimilarity*: two (or more) groups are similarly distributed whenever "the overall populations of the two groups take the same values with the same frequency." [Gini, 1914].
- When does a set of distributions display more dissimilarity than another?

Motivation

- We analyze the inequalities between distributions (*groups*) of an ordered attribute.
- *Dissimilarity*: two (or more) groups are similarly distributed whenever "the overall populations of the two groups take the same values with the same frequency." [Gini, 1914].
- When does a set of distributions display more dissimilarity than another?
- Relevant question for:
  - *Unfair inequality* [Fleurbaye 2008, Roemer Trannoy 2016, Ferreira and Peragine 2018, ...]
  - *Discrimination* [Gastwirth 1975, Dagum 1980, Jenkins 1994, Le Breton et al 2012]
  - *Mobility* [Dardanoni 1993, Van de gaer et al 2001, Jantti and Jenkins 2015]
  - *Distance between distributions* [Shorrocks 1982, Ebert 1984, Magdalou and Nock 2011]

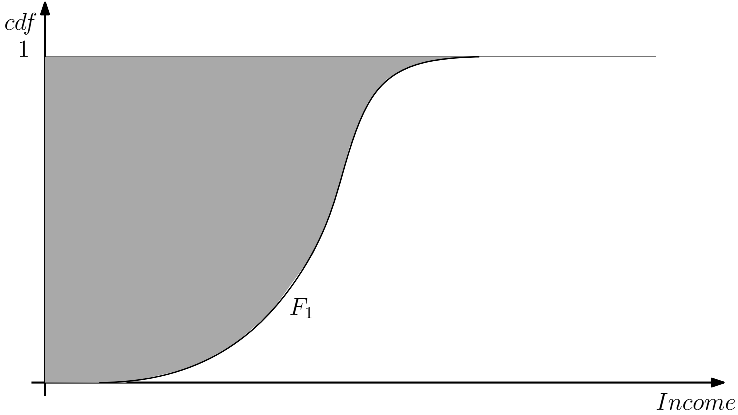
Motivation

- Consider two distributions of a cardinal attribute.
- A basic criterion is the *difference in averages*.
-
-



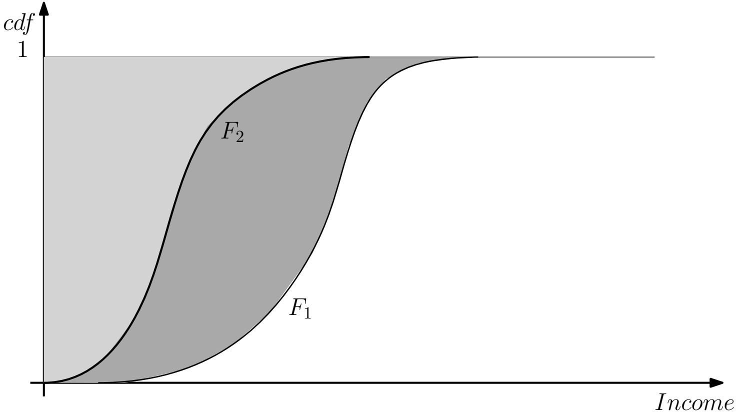
Motivation

- Consider two distributions of a cardinal attribute.
- A basic criterion is the *difference in averages*.
-
-



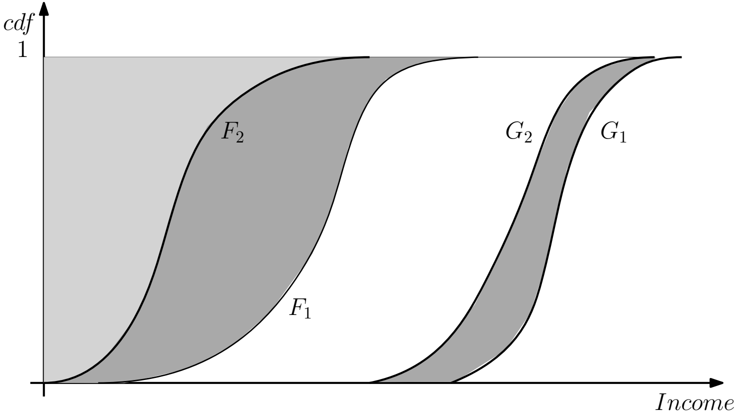
Motivation

- Consider two distributions of a cardinal attribute.
- A basic criterion is the *difference in averages*.
-
-



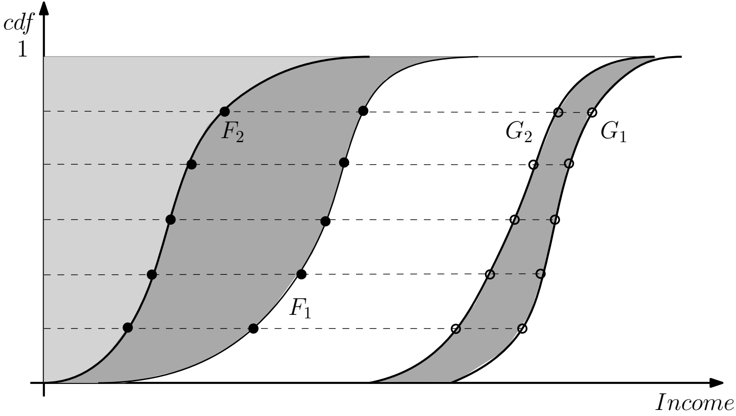
Motivation

- Consider two distributions of a cardinal attribute.
- A basic measure criterion is the *difference in averages*.
- The criterion is useful to compare situations G_1, G_2 versus F_1, F_2.
-



Motivation

- Consider two distributions of a cardinal attribute.
- A basic measure criterion is the *difference in averages*.
- The criterion is useful to compare situations G_1, G_2 versus F_1, F_2.
- A more robust approach is *gap curve dominance* [Andreoli et al, 2019]



Motivation

- These and similar criteria are **translation invariant** and **robust**, but...
- ... a desirable criterion should also be **scale invariant** and, in a broader sense, **invariant to monotone transformations** of the data.
  - Model specification (FE, trends,...),
  - Choice of scale ($, ranks, log$)
  - ...,

Motivation

- These and similar criteria are **translation invariant** and **robust**, but...
- ... a desirable criterion should also be **scale invariant** and, in a broader sense, **invariant to monotone transformations** of the data.
  - Model specification (FE, trends,...),
  - Choice of scale ($, ranks, log$)
  - ...,
- If the criterion preserves **only ordinal information** it is also useful for studying:
  - Self-assessed health.
  - Skills.
  - Composite indicators of well-being.
  - Ordered alternatives (jobs, neighborhoods, schools).
  - ...

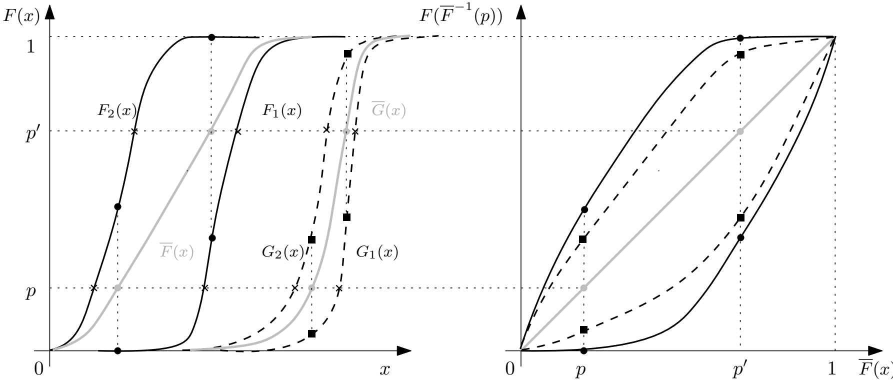
Motivation

- These and similar criteria are **translation invariant** and **robust**, but...
- ... a desirable criterion should also be **scale invariant** and, in a broader sense, **invariant to monotone transformations** of the data.
  - Model specification (FE, trends,...),
  - Choice of scale ($, ranks, log$)
  - ...,
- If the criterion preserves **only ordinal information** it is also useful for studying:
  - Self-assessed health.
  - Skills.
  - Composite indicators of well-being.
  - Ordered alternatives (jobs, neighborhoods, schools).
  - ...

Contribution

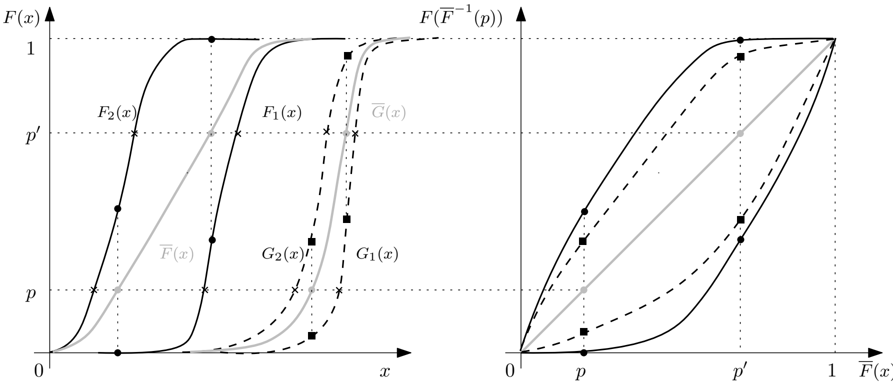
We develop a dissimilarity criterion for comparing distributions $F_1, \ldots, F_d$ to $G_1, \ldots, G_d$ that is **robust**, **invariant to monotone transformations** *and hence* **preserves ordinal information**. *Our main result offers an* **axiomatic derivation** *of the criterion.*

Motivation



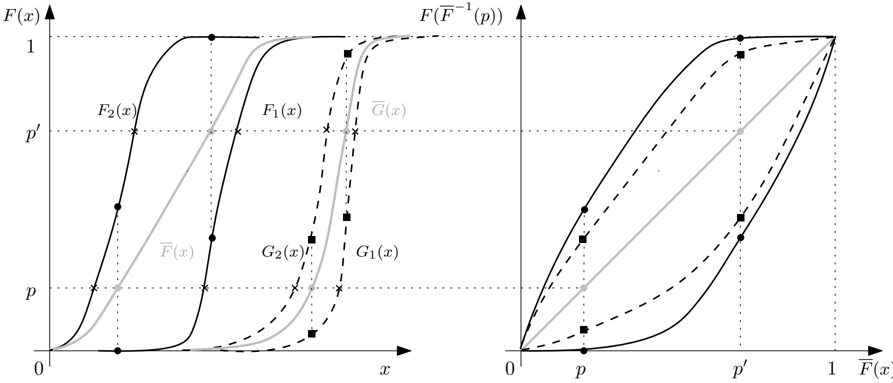
- Choose reference distributions (endogenous) $\overline{F}$ and $\overline{G}$ (in gray).

Motivation



- Choose reference distributions (endogenous) $\overline{F}$ and $\overline{G}$ (in gray).
- Measure proportions $F_1(.)$, $F_2(.)$, $G_1(.)$, $G_2()$ at quantiles $\overline{F}^{-1}(p)$ and $\overline{G}^{-1}(p)$ at proportions $p, p', \ldots$ (● and ■ symbols)

Motivation



- Choose reference distributions (endogenous) $\overline{F}$ and $\overline{G}$ (in gray).
- Measure proportions $F_1(.)$, $F_2(.)$, $G_1(.)$, $G_2()$ at quantiles $\overline{F}^{-1}(p)$ and $\overline{G}^{-1}(p)$ at proportions $p, p', \ldots$ (● and ■ symbols)
- **Criterion:**

$$|G_1(\overline{G}^{-1}(p)) - G_2(\overline{G}^{-1}(p))| \leq |F_1(\overline{F}^{-1}(p)) - F_2(\overline{F}^{-1}(p))|, \ \forall p \in [0, 1]$$

Outline of the presentation

- **A dissimilarity criterion for discrete (empirical) distributions**:
  - matrix notation,
  - piecewise linear representations.
- **Axiomatic model**;
- **Characterization**;
- **Additional results**
  - Dissimilarity indices
  - Implementable conditions
  - Empirical comparisons
  - Dissimilarity, discrimination and distance between distributions.
- **Empirical illustration**: Unfair inequality and education reforms in Sweden [Meghir and Palme 2005]:

Notation

- $\mathbf{A} = [a_{ij}]_{j=1,\ldots,n}^{i=1,\ldots,d} \in \mathcal{M}_d$ is a **distribution matrix**
  - d groups by n classes.
  - a_{ij} is the proportion of group i observed in class j.
  - Matrices in $\mathcal{M}_d$ have fixed d but variable n.
- $\overrightarrow{\mathbf{A}} \in \mathbb{R}_{+}^{d,n_A}$ is a **cumulative distribution matrix**
  - $\overrightarrow{\mathbf{a}}_k := \sum_{j=1}^{k} \mathbf{a}_j$.
- h **first order stochastic dominates** that of groups ℓ whenever $\overrightarrow{a}_{hj} \leq \overrightarrow{a}_{\ell j}$ for all $j = 1, \ldots, n$, with a strict inequality ($<$) holding for at least a class.

Notation

- $p_j \in [0,1]$ is the **average cumulative distributions across groups** in j.
  - $p_j = \frac{1}{d}\sum_i \overrightarrow{a}_{ij} \in [0,1]$.
- $\overrightarrow{a}_i(p) \in [0,1]$ is the **cumulative group distribution**
  - onto function specific of each group i
  - $\overrightarrow{a}_i(p_j) = \overrightarrow{a}_{ij}$
  - $\overrightarrow{a}_i(0) = 0$
  - $\overrightarrow{a}_i(p_n) = 1$
  - For $p \in (p_{j-1}, p_j)$ it solves $p = \frac{1}{d}\sum_i \overrightarrow{a}_i(p)$, obtained by linear interpolation of $\overrightarrow{a}_{ij}$ and $\overrightarrow{a}_{ij+1}$:

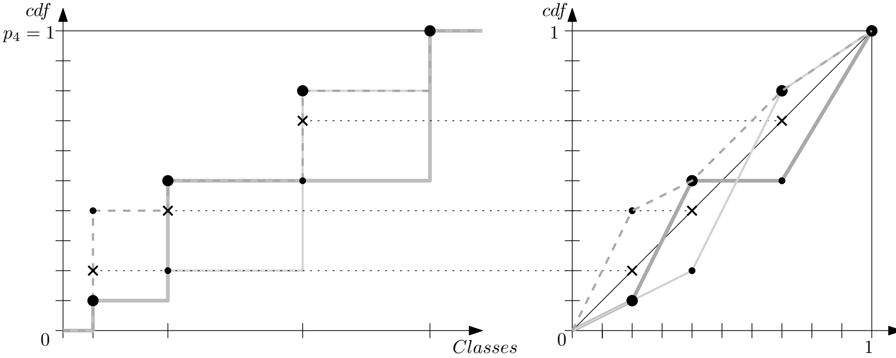
$$\overrightarrow{\mathbf{a}}(p) := (\overrightarrow{a}_1(p), \ldots, \overrightarrow{a}_d(p))^t = \overrightarrow{\mathbf{a}}_{j-1} + \frac{p - p_{j-1}}{p_j - p_{j-1}}\mathbf{a}_j.$$

- Plotting $\overrightarrow{a}_i(p)$ across levels $p \in [0,1]$ gives instead a **piecewise linear graph** on the unit interval domain

Notation

Example

$$\mathbf{A} = \begin{pmatrix} 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.4 & 0 & 0.5 \\ 0.1 & 0.1 & 0.6 & 0.2 \end{pmatrix} \quad \textit{and} \quad \overrightarrow{\mathbf{A}} = \begin{pmatrix} 0.4 & 0.5 & 0.8 & 1 \\ 0.1 & 0.5 & 0.5 & 1 \\ 0.1 & 0.2 & 0.8 & 1 \end{pmatrix}. \quad (1)$$



The dissimilarity criterion

Definition

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$, $\mathbf{B}$ is at most as dissimilar as $\mathbf{A}$, which we denote $\mathbf{B} \preccurlyeq^D \mathbf{A}$ if and only if for all $p \in [0,1]$

$$\sum_i^h \overrightarrow{b}_{(i)}(p) \geq \sum_i^h \overrightarrow{a}_{(i)}(p), \quad h = 1, \ldots, d. \tag{2}$$

We say that $\mathbf{B}$ is **as most as dissimilar as $\mathbf{A}$** if the proportions of the groups adding up to the bottom $p100\%$ of the average of the cumulative distributions across groups in $\mathbf{B}$ (i.e $\overrightarrow{\mathbf{b}}(p)$) are unambiguously less dispersed than the corresponding proportions in $\mathbf{A}$ (i.e. $\overrightarrow{\mathbf{a}}(p)$), for any $p \in [0,1]$.

The dissimilarity criterion

Definition

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$, $\mathbf{B}$ is at most as dissimilar as $\mathbf{A}$, which we denote $\mathbf{B} \preccurlyeq^D \mathbf{A}$ if and only if for all $p \in [0, 1]$

$$\sum_{i}^{h} \overrightarrow{b}_{(i)}(p) \geq \sum_{i}^{h} \overrightarrow{a}_{(i)}(p), \quad h = 1, \ldots, d. \tag{2}$$

We say that $\mathbf{B}$ is **as most as dissimilar as** $\mathbf{A}$ if the proportions of the groups adding up to the bottom $p100\%$ of the average of the cumulative distributions across groups in $\mathbf{B}$ (i.e $\overrightarrow{\mathbf{b}}(p)$) are unambiguously less dispersed than the corresponding proportions in $\mathbf{A}$ (i.e. $\overrightarrow{\mathbf{a}}(p)$), for any $p \in [0, 1]$.

Remark

When $d = 2$, the Lorenz dominance condition (2) can be equivalently stated as

$$|\overrightarrow{b}_1(p) - \overrightarrow{b}_2(p)| \leq |\overrightarrow{a}_1(p) - \overrightarrow{a}_2(p)|, \quad \forall p \in [0, 1].$$

The dissimilarity criterion

The dissimilarity criterion $\preccurlyeq^D$ sets out a **partial order** of distribution matrices:

- Lorenz dominance of shares at any p,
- For all p.

The dissimilarity criterion

The dissimilarity criterion $\preccurlyeq^D$ sets out a **partial order** of distribution matrices:

- Lorenz dominance of shares at any p,
- For all p.

1) Transitivity.

Remark

For any $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{M}_d$*, if* $\mathbf{B} \preccurlyeq^D \mathbf{A}$ *and* $\mathbf{C} \preccurlyeq^D \mathbf{B}$ *then* $\mathbf{C} \preccurlyeq^D \mathbf{A}$*.*

The dissimilarity criterion

The dissimilarity criterion $\preccurlyeq^D$ sets out a **partial order** of distribution matrices:

- Lorenz dominance of shares at any p,
- For all p.

1) Transitivity.

Remark

For any $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{M}_d$*, if* $\mathbf{B} \preccurlyeq^D \mathbf{A}$ *and* $\mathbf{C} \preccurlyeq^D \mathbf{B}$ *then* $\mathbf{C} \preccurlyeq^D \mathbf{A}$*.*

2) Boundedness. Let the **perfect similarity S** and **maximal dissimilarity D** matrices:

$$\mathbf{S} := \begin{pmatrix} \mathbf{s}' \\ \vdots \\ \mathbf{s}' \end{pmatrix} \quad \text{and} \quad \mathbf{D} := \begin{pmatrix} \mathbf{d}'_1 & \dots & \mathbf{0}'_{n_d} \\ \vdots & \ddots & \vdots \\ \mathbf{0}'_{n_1} & \dots & \mathbf{d}'_d \end{pmatrix}. \tag{3}$$

Remark

For any $\mathbf{S}, \mathbf{D}, \mathbf{A} \in \mathcal{M}_d$ *where* $\mathbf{S}$ *and* $\mathbf{D}$ *are as in (3),* $\mathbf{S} \preccurlyeq^D \mathbf{A} \preccurlyeq^D \mathbf{D}$*.*

Remark

Let $\mathbf{S}, \mathbf{S}'$ *be two distinct perfect similarity matrices and* $\mathbf{D}, \mathbf{D}'$ *be two distinct maximal dissimilarity matrices, then* $\mathbf{S} \sim^D \mathbf{S}'$ *and* $\mathbf{D} \sim^D \mathbf{D}'$*.*

Axiomatic model

A *dissimilarity ordering* is a complete and transitive binary relation $\preccurlyeq$ on the set $\mathcal{M}_d$ with symmetric part $\sim$, that ranks $\mathbf{B} \preccurlyeq \mathbf{A}$ whenever $\mathbf{B}$ is *at most as dissimilar as* $\mathbf{A}$.

Axiomatic model

A *dissimilarity ordering* is a complete and transitive binary relation $\preccurlyeq$ on the set $\mathcal{M}_d$ with symmetric part $\sim$, that ranks $\mathbf{B} \preccurlyeq \mathbf{A}$ whenever $\mathbf{B}$ is *at most as dissimilar as* $\mathbf{A}$.

Axiom

E (Exchange) *For any* $\mathbf{A}$, $\mathbf{B} \in \mathcal{M}_d$ *with* $n_A = n_B = n$ *where group* h *dominates group* ℓ *and* $k' > k$, *if* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by an exchange transformation such that (i)* $b_{hk} = a_{hk} + \varepsilon$ *and* $b_{hk'} = a_{hk'} - \varepsilon$, *(ii)* $b_{\ell k} = a_{\ell k} - \varepsilon$ *and* $b_{\ell k'} = a_{\ell k'} + \varepsilon$, *(iii)* $b_{ij} = a_{ij}$ *in all other cases, (iv)* $\varepsilon > 0$ *so that if* $\overrightarrow{a}_{ij} \leq \overrightarrow{a}_{i'j}$ *then* $\overrightarrow{b}_{ij} \leq \overrightarrow{b}_{i'j}$ *for all groups* $i \neq i'$ *and for all classes* j, *then* $\mathbf{B} \preccurlyeq \mathbf{A}$.

Example

$$\mathbf{B} = \begin{pmatrix} 0.4 & 0.1 & 0.3 - \varepsilon & 0.2 + \varepsilon \\ 0.1 & 0.4 & 0 + \varepsilon & 0.5 - \varepsilon \\ 0.1 & 0.1 & 0.6 & 0.2 \end{pmatrix} \preccurlyeq \mathbf{A}. \tag{4}$$

Axiomatic model

Next set of axioms allows to modify the shape of a distribution matrix but does not affect its informational content. Cardinality concerns are lost.

Axiomatic model

Next set of axioms allows to modify the shape of a distribution matrix but does not affect its informational content. Cardinality concerns are lost.

Axiom

IEC (Independence from Empty Classes) *For any* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D} \in \mathcal{M}_d$ *and* $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$, *if* $\mathbf{B} = (\mathbf{A}_1, \mathbf{0}_d, \mathbf{A}_2)$, $\mathbf{C} = (\mathbf{0}_d, \mathbf{A})$, $\mathbf{D} = (\mathbf{A}, \mathbf{0}_d)$ *then* $\mathbf{B} \sim \mathbf{C} \sim \mathbf{D} \sim \mathbf{A}$.

Axiomatic model

Next set of axioms allows to modify the shape of a distribution matrix but does not affect its informational content. Cardinality concerns are lost.

Axiom

IEC (Independence from Empty Classes) *For any* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D} \in \mathcal{M}_d$ *and* $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$, *if* $\mathbf{B} = (\mathbf{A}_1, \mathbf{0}_d, \mathbf{A}_2)$, $\mathbf{C} = (\mathbf{0}_d, \mathbf{A})$, $\mathbf{D} = (\mathbf{A}, \mathbf{0}_d)$ *then* $\mathbf{B} \sim \mathbf{C} \sim \mathbf{D} \sim \mathbf{A}$.

Axiom

ISC (Independence from Split of Classes) *For any* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *with* $n_B = n_A + 1$, *if* $\exists j$ *such that* $\mathbf{b}_j = \beta \mathbf{a}_j$ *and* $\mathbf{b}_{j+1} = (1-\beta)\mathbf{a}_j$ *with* $\beta \in (0,1)$, *while* $\mathbf{b}_k = \mathbf{a}_k$ $\forall k < j$ *and* $\mathbf{b}_{k+1} = \mathbf{a}_k$ $\forall k > j$, *then* $\mathbf{B} \sim \mathbf{A}$.

Axiomatic model

Next set of axioms allows to modify the shape of a distribution matrix but does not affect its informational content. Cardinality concerns are lost.

Axiom

IEC (Independence from Empty Classes) *For any* $\mathbf{A},\ \mathbf{B},\ \mathbf{C},\ \mathbf{D} \in \mathcal{M}_d$ *and* $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$, *if* $\mathbf{B} = (\mathbf{A}_1, \mathbf{0}_d, \mathbf{A}_2)$, $\mathbf{C} = (\mathbf{0}_d, \mathbf{A})$, $\mathbf{D} = (\mathbf{A}, \mathbf{0}_d)$ *then* $\mathbf{B} \sim \mathbf{C} \sim \mathbf{D} \sim \mathbf{A}$.

Axiom

ISC (Independence from Split of Classes) *For any* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *with* $n_B = n_A + 1$, *if* $\exists j$ *such that* $\mathbf{b}_j = \beta \mathbf{a}_j$ *and* $\mathbf{b}_{j+1} = (1-\beta)\mathbf{a}_j$ *with* $\beta \in (0,1)$, *while* $\mathbf{b}_k = \mathbf{a}_k\ \forall k < j$ *and* $\mathbf{b}_{k+1} = \mathbf{a}_k\ \forall k > j$, *then* $\mathbf{B} \sim \mathbf{A}$.

Example

$$\begin{pmatrix} 0.4 & \beta 0.1 & (1-\beta)0.1 & 0.3 & 0.2 \\ 0.1 & \beta 0.4 & (1-\beta)0.4 & 0 & 0.5 \\ 0.1 & \beta 0.1 & (1-\beta)0.1 & 0.6 & 0.2 \end{pmatrix} \backsim \begin{pmatrix} 0 & 0.4 & 0.1 & 0 & 0.3 & 0.2 & 0 \\ 0 & 0.1 & 0.4 & 0 & 0 & 0.5 & 0 \\ 0 & 0.1 & 0.1 & 0 & 0.6 & 0.2 & 0 \end{pmatrix} \backsim \mathbf{A} \quad (5)$$

Axiomatic model

Next axioms emphasize the extent of disadvantage of one group over another rather than the sign of it.

Axiomatic model

Next axioms emphasize the extent of disadvantage of one group over another rather than the sign of it.

Axiom

IPG (Independence from Permutations of Groups) *For any* $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$, *if* $\mathbf{B} = \mathbf{\Pi}_d \cdot \mathbf{A}$ *for a permutation matrix* $\mathbf{\Pi}_d \in \mathcal{P}_d$ *then* $\mathbf{B} \sim \mathbf{A}$.

Axiomatic model

Next axioms emphasize the extent of disadvantage of one group over another rather than the sign of it.

Axiom

IPG (Independence from Permutations of Groups) *For any* $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$*, if* $\mathbf{B} = \mathbf{\Pi}_d \cdot \mathbf{A}$ *for a permutation matrix* $\mathbf{\Pi}_d \in \mathcal{P}_d$ *then* $\mathbf{B} \sim \mathbf{A}$*.*

Axiom

I (Interchange of Groups) *For any* $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ *with* $n_A = n_B = n$*, if* $\exists \mathbf{\Pi}_{h,\ell} \in \mathcal{P}_d$ *permuting only groups* h *and* ℓ *whenever* $\overrightarrow{a}_{hk} = \overrightarrow{a}_{\ell k}$*, such that* $\mathbf{B} = (\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{\Pi}_{h,\ell} \cdot \mathbf{a}_{k+1}, \ldots, \mathbf{\Pi}_{h,\ell} \cdot \mathbf{a}_{n_A})$*, then* $\mathbf{B} \sim \mathbf{A}$*.*

Axiomatic model

Next axioms emphasize the extent of disadvantage of one group over another rather than the sign of it.

Axiom

IPG (Independence from Permutations of Groups) *For any* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$, *if* $\mathbf{B} = \mathbf{\Pi}_d \cdot \mathbf{A}$ *for a permutation matrix* $\mathbf{\Pi}_d \in \mathcal{P}_d$ *then* $\mathbf{B} \sim \mathbf{A}$.

Axiom

I (Interchange of Groups) *For any* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *with* $n_A = n_B = n$, *if* $\exists \mathbf{\Pi}_{h,\ell} \in \mathcal{P}_d$ *permuting only groups* h *and* ℓ *whenever* $\vec{a}_{hk} = \vec{a}_{\ell k}$, *such that* $\mathbf{B} = (\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{\Pi}_{h,\ell} \cdot \mathbf{a}_{k+1}, \dots, \mathbf{\Pi}_{h,\ell} \cdot \mathbf{a}_{n_A})$, *then* $\mathbf{B} \sim \mathbf{A}$.

Example

$$\begin{pmatrix} 0.1 & 0.4 & 0 & 0.5 \\ 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.6 & 0.2 \end{pmatrix} \backsim \begin{pmatrix} 0.4 & 0.1 & 0 & 0.5 \\ 0.1 & 0.4 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.6 & 0.2 \end{pmatrix} \backsim \mathbf{A} \qquad (6)$$

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B} \ \preccurlyeq\ \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^D \mathbf{A}$.

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B} \ \preccurlyeq\ \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^D \mathbf{A}$.

- Axiom E combines concerns for **distance and correlation reduction** [Epstain Tanny 1980, Tchen 1980, Atkinson Bourguignon 1982] as two equivalent perspectives. Axiom I eliminates concerns for correlation.

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B} \preccurlyeq \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^D \mathbf{A}$.

- Axiom E combines concerns for **distance and correlation reduction** [Epstain Tanny 1980, Tchen 1980, Atkinson Bourguignon 1982] as two equivalent perspectives. Axiom I eliminates concerns for correlation.
- When $\mathcal{M}_d$ are **income mobility matrices**, the Theorem offer and alternative characterization of the **orthant order** [Dardanoni 1993, Jantti and Jenkins 2015]

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B}\ \preccurlyeq\ \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^D \mathbf{A}$.

- Axiom E combines concerns for **distance and correlation reduction** [Epstain Tanny 1980, Tchen 1980, Atkinson Bourguignon 1982] as two equivalent perspectives. Axiom I eliminates concerns for correlation.
- When $\mathcal{M}_d$ are **income mobility matrices**, the Theorem offer and alternative characterization of the **orthant order** [Dardanoni 1993, Jantti and Jenkins 2015]
- When $\mathcal{M}_d$ represent **intergenerational mobility matrices**, $\preccurlyeq^D$ regards **origin independence** as a reference [Shorrocks 1978, Kanbur Stiglitz 2016].

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B}\ \preccurlyeq\ \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^{D} \mathbf{A}$.

- Axiom E combines concerns for **distance and correlation reduction** [Epstain Tanny 1980, Tchen 1980, Atkinson Bourguignon 1982] as two equivalent perspectives. Axiom I eliminates concerns for correlation.
- When $\mathcal{M}_d$ are **income mobility matrices**, the Theorem offer and alternative characterization of the **orthant order** [Dardanoni 1993, Jantti and Jenkins 2015]
- When $\mathcal{M}_d$ represent **intergenerational mobility matrices**, $\preccurlyeq^{D}$ regards **origin independence** as a reference [Shorrocks 1978, Kanbur Stiglitz 2016].
- The Theorem extends the (intergenerational) mobility orders to matrices that are non-monotone with different margins. Useful to in **inequality of opportunity IOP analysis** [Roemer and Trannoy 2016, Andreoli et al 2019, Andreoli et al 2022]

Main result /1

The intersection of the dissimilarity orderings $\preccurlyeq$ (which is a partial order [see Donaldson and Weymark 1998]) characterizes the dissimilarity criterion

Theorem

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *the following statements are equivalent:*

(i) $\mathbf{B} \preccurlyeq \mathbf{A}$ *for every ordering* $\preccurlyeq$ *satisfying axioms E, SC, IEC, IPG and I,*

(ii) $\mathbf{B} \preccurlyeq^D \mathbf{A}$.

- Axiom E combines concerns for **distance and correlation reduction** [Epstain Tanny 1980, Tchen 1980, Atkinson Bourguignon 1982] as two equivalent perspectives. Axiom I eliminates concerns for correlation.
- When $\mathcal{M}_d$ are **income mobility matrices**, the Theorem offer and alternative characterization of the **orthant order** [Dardanoni 1993, Jantti and Jenkins 2015]
- When $\mathcal{M}_d$ represent **intergenerational mobility matrices**, $\preccurlyeq^D$ regards **origin independence** as a reference [Shorrocks 1978, Kanbur Stiglitz 2016].
- The Theorem extends the (intergenerational) mobility orders to matrices that are non-monotone with different margins. Useful to in **inequality of opportunity IOP analysis** [Roemer and Trannoy 2016, Andreoli et al 2019, Andreoli et al 2022]
- The Theorem yields a normative justification to robust IOP comparisons with **ordinal variables** [Ferreira Gignoux 2011].

Main result /2

- The $\preccurlyeq^D$ gives rise to an **orthant test** [Ch. 6 in Shaked Shantikuman 2006].
- Axiom E invokes a stronger but appealing principle (that groups are ordered by SD) than [Tchen 1980]. Axiom E weaker than operations that characterize the supermodularity order [Mayer and Strulovici 2013].

Main result /2

- The $\preccurlyeq^D$ gives rise to an **orthant test** [Ch. 6 in Shaked Shantikuman 2006].
- Axiom E invokes a stronger but appealing principle (that groups are ordered by SD) than [Tchen 1980]. Axiom E weaker than operations that characterize the supermodularity order [Mayer and Strulovici 2013].
- Axiom E is incompatible with **Merge operations**, regarded as unambiguously dissimilarity-reducing when **outcomes are categories** [Andreoli and Zoli, SCW 2022].
  - Merge operations characterize **Matrix Majorization**,
  - Weakened to **zonotope inclusion criterion**.

Example

Consider merging (element by element) classes 2 and 3 of matrix $\tilde{\mathbf{A}}$ *and then splitting in proportion* $5/8$*. This gives:*

$$\tilde{\mathbf{A}} \rightarrow \begin{pmatrix} 0.4 & 0 & 0.4 & 0.2 \\ 0.1 & 0 & 0.4 & 0.5 \end{pmatrix} \rightarrow \begin{pmatrix} 0.4 & 0.4\frac{5}{8} & 0.4\frac{3}{8} & 0.2 \\ 0.1 & 0.4\frac{5}{8} & 0.4\frac{3}{8} & 0.5 \end{pmatrix} = \tilde{\mathbf{B}}$$

$\tilde{\mathbf{B}}$ *unambiguously less dissimilar than* $\tilde{\mathbf{A}}$ *[Andreoli and Zoli, SCW 2022].*
But, *for* $\varepsilon = 0.15$*:*

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0.4 & 0.25-\varepsilon & 0.15+\varepsilon & 0.2 \\ 0.1 & 0.25+\varepsilon & 0.15-\varepsilon & 0.5 \end{pmatrix} \preccurlyeq^D \tilde{\mathbf{B}}!!!!$$

Useful facts

- Sketch of the proof. ▸ Go!
- Characterization of a family of dissimilarity indices. ▸ Go!
- Implementation of the dissimilarity criterion. ▸ Go!
- Empirical dissimilarity criterion. ▸ Go!
- The geometry of dissimilarity. ▸ Go!
- Dissimilarity, discrimination and distance. ▸ Go!

Application: Evaluating the Swedish education reform

The **Swedish education reform**:

- increased compulsory education duration, abolished streaming after grade six and introduced a uniform national curriculum.
- gradually introduced across Swedish municipalities in 1949 until 1962.
- Quasi-random variation across space and cohorts [Meghir Palme 2005].
- Significant average effects on earnings [Meghir Palme 2005, Fisher et al 2020] education [Holmlund 2007], mortality [Lager Torssander 2012], health [Meghir et at 2018].

Objective: assess the consequences of the reform on unfair inequality in income, along circumstances of birth.

Application: Evaluating the Swedish education reform

Sample design:

- Cohorts 1948 (pre-reform) and 1953 (post-reform)
- About 18,000 boys and girls, whose income is observed 1985 though 1996.
- 65% of sample lives in municipalities that switch into the reformed system about 1962.

Application: Evaluating the Swedish education reform

Sample design:

- Cohorts 1948 (pre-reform) and 1953 (post-reform)
- About 18,000 boys and girls, whose income is observed 1985 though 1996.
- 65% of sample lives in municipalities that switch into the reformed system about 1962.

Goal: compare income opportunities of post-reform cohort in treated municipalities (**treatment**) with the **control** municipalities.

Application: Evaluating the Swedish education reform

Sample design:

- Cohorts 1948 (pre-reform) and 1953 (post-reform)
- About 18,000 boys and girls, whose income is observed 1985 though 1996.
- 65% of sample lives in municipalities that switch into the reformed system about 1962.

Goal: compare income opportunities of post-reform cohort in treated municipalities (**treatment**) with the **control** municipalities.

Identification: Quasi random assignment of municipalities switching reform status, clear of cohort, municipality, county, year FE and trends.

Application: Evaluating the Swedish education reform

Sample design:

- Cohorts 1948 (pre-reform) and 1953 (post-reform)
- About 18,000 boys and girls, whose income is observed 1985 though 1996.
- 65% of sample lives in municipalities that switch into the reformed system about 1962.

Goal: compare income opportunities of post-reform cohort in treated municipalities (**treatment**) with the **control** municipalities.

Identification: Quasi random assignment of municipalities switching reform status, clear of cohort, municipality, county, year FE and trends.

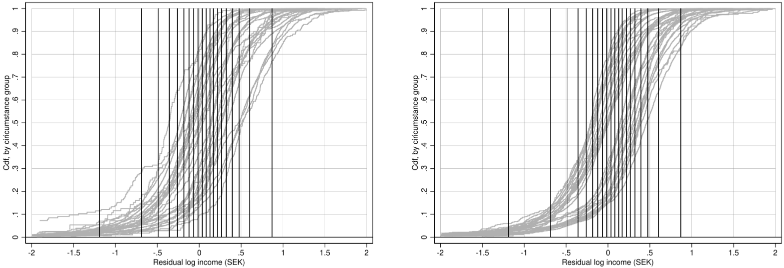
Estimating model: Let y_{itcmd} be the log-income observed in year t for children i when aged about 40 years old, born in cohort c and municipality m and living in a treatment ($d = 1$) or control ($d = 0$) municipality. The log-income process is decomposed according to the following specification:

$$y_{itcmd} = \theta_0 + \theta_t + \theta_c + \theta_m + \theta_d + \theta_{tc} + \theta_{td} + \gamma_0 t + \gamma_c t + \gamma_m t + \varepsilon_{itcmd}, \quad (7)$$

Predict residuals $\hat{\varepsilon_{itcmd}}$ and distinguish along the lines of **treatment vs control** groups and **gender, ability, location and parental background characteristics** (**$d = 32$ groups**).

Application: Evaluating the Swedish education reform

The specification of the model residuals may affect evaluations that retain cardinal information (vertical bars are vingitiles):

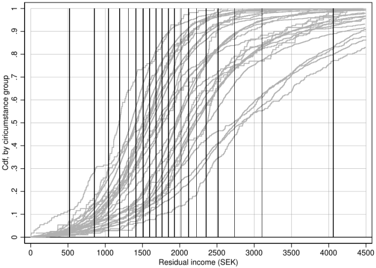


(a) Control group

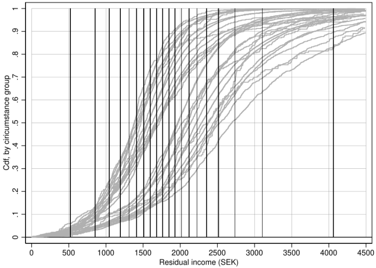
(b) Treatment group

Application: Evaluating the Swedish education reform

The specification of the model residuals may affect evaluations that retain cardinal information (vertical bars are vingitiles):



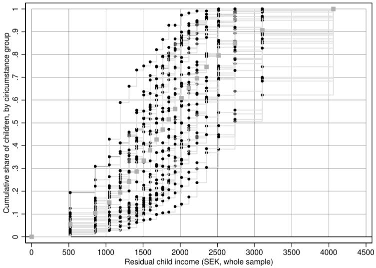
(c) Control group



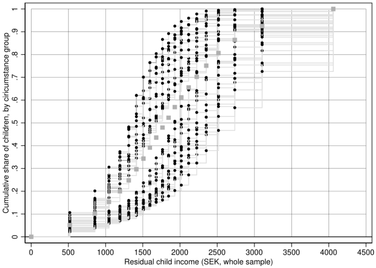
(d) Treatment group

Application: Evaluating the Swedish education reform

For ease of exposition, comparisons are limited to **earnings vingitiles**, yielding two distribution matrices of size 32×20 for the treatment (**T**) and control (**C**) cases.



(e) Control group **C**

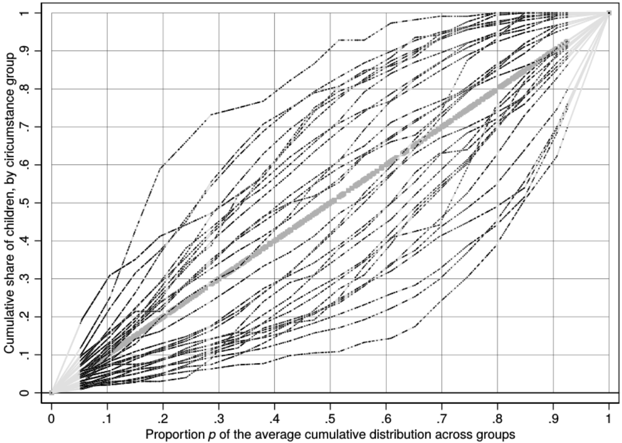


(f) Treatment group **T**

Application: Evaluating the Swedish education reform

From **T** and **C** we derive the empirical representations of the cumulative group distributions $\overrightarrow{\mathbf{t}}(p)$ and $\overrightarrow{\mathbf{c}}(p)$ for $p \in [0, 1]$, respectively (in gray).

We also plot the (finite) set of points for which it is sufficient to test Lorenz dominance (at fixed p) in order to conclude on the null $\mathbf{T} \preccurlyeq^D \mathbf{C}$ (**in black**)

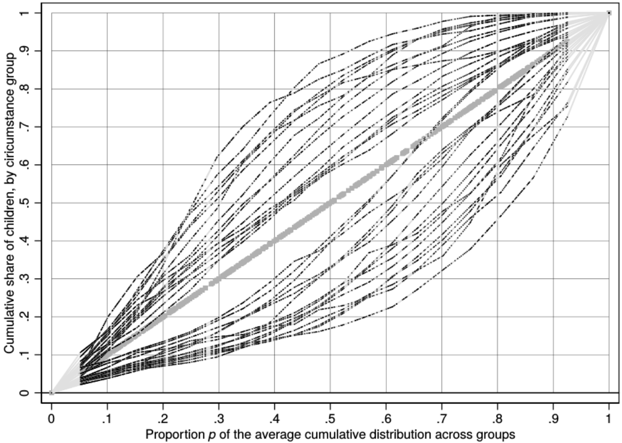


(g) Control group $\overrightarrow{\mathbf{c}}(p)$

Application: Evaluating the Swedish education reform

From **T** and **C** we derive the empirical representations of the cumulative group distributions $\overrightarrow{\mathbf{t}}(p)$ and $\overrightarrow{\mathbf{c}}(p)$ for $p \in [0, 1]$, respectively (in gray).

We also plot the (finite) set of points for which it is sufficient to test Lorenz dominance (at fixed p) in order to conclude on the null $\mathbf{T} \preccurlyeq^{D} \mathbf{C}$ (**in black**)



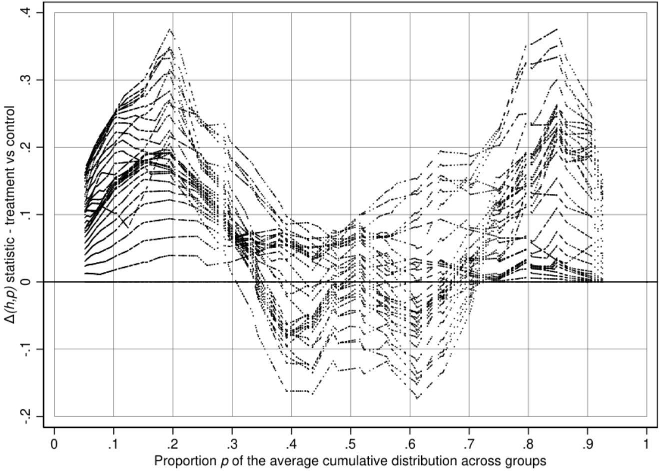
(h) Treatment group $\overrightarrow{\mathbf{t}}(p)$

Application: Evaluating the Swedish education reform

Dominance test for the null $\mathbf{T} \preccurlyeq^{D} \mathbf{C}$ builds on the statistic

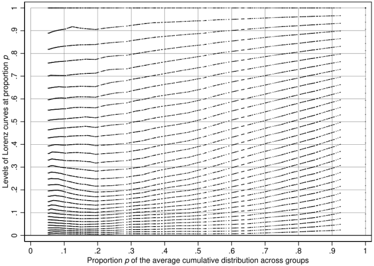
$$\sum_{i}^{h} \overrightarrow{b}_{(i)}(p_j) \geq \sum_{i}^{h} \overrightarrow{a}_{(i)}(p_j), \quad h = 1, \ldots, d,$$

issued at a finite number of intercepts $p_1, \ldots, p_{493}$ ▸ Implementation

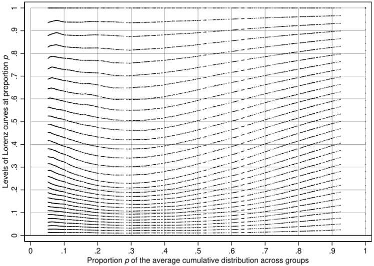


Application: Evaluating the Swedish education reform

The test is not informative about where inequalities in groups distributions are stronger **over the domain of** p. A relative version of the test allows to deal with this issue, considering first the groups cumulative distributions relative to the average, dp, and then constructing the **Lorenz curves coordinates** as follows, for $p_1, \ldots, p_{493}$:



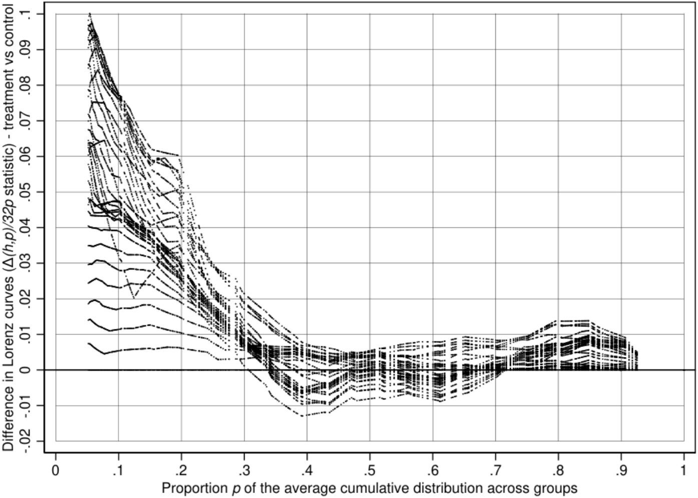
(i) Control group $\sum_{i=1}^{h} \frac{1}{dp_j} \overrightarrow{c}^{*}_{(i)j}$



(j) Treatment group $\sum_{i=1}^{h} \frac{1}{dp_j} \overrightarrow{c}^{*}_{(i)j}$

Application: Evaluating the Swedish education reform

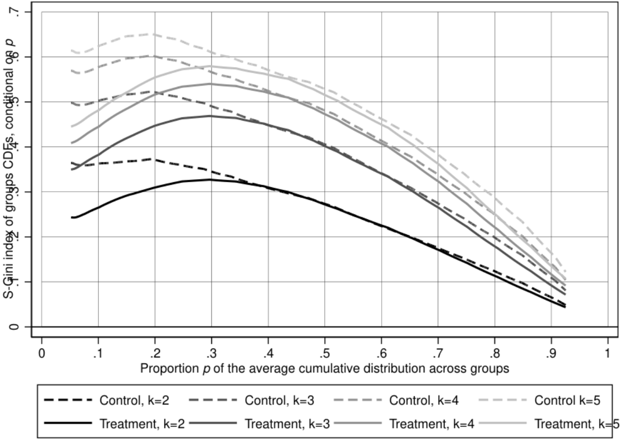
Using the relative statistics based on differences in Lorenz curves at intercepts $p_1, \ldots, p_{493}$, we conclude that the education reform has made income opportunities more equal or low-income achievers, whereas the effects are ambiguous in the middle of the distribution. ▸ Implementation



Application: Evaluating the Swedish education reform

Such differences my be statistically significant, but they are more than compensated by improvements at the bottom according to aggregate measures of dissimilarity.

The figure reports the estimator for $\sum_{i=1}^{32} w_i t_{(i)}(p_j)$ and $\sum_{i=1}^{32} w_i c_{(i)}(p_j)$ for selected $p \in [0,1]$, where w_i is the S-Gini weighting function (parametrized by $k = 1, \ldots, 5$). ▸ Indices



Conclusions

A new **dissimilarity criterion** $\preccurlyeq^D$ is introduced:

- Multi-group ($d \geq 2$),
- Preserves ordinal information,
- Robust,
- Partial order.

Axiomatic characterization of $\preccurlyeq^D$:

- Intersection of dissimilarity orderings,
- Based on simple operations,
- Exchange: widely used,
- Interchange: introduce concerns for distance of distributions,
- Operations break down more effects of more complex transformations (eg. policy)

A **policy evaluation** exercise using $\preccurlyeq^D$:

- Evaluate the distributional impact of the Swedish reform on unfair inequality,
- Test is rejected,
- Yet, violations are mild and concentrated in the middle.
- Aggregate measures support reduction of unfair inequality.

Implementation

Back

Assessing $\mathbf{B} \preccurlyeq^D \mathbf{A}$ requires and infinity of Lorenz curve dominance comparisons.

For a special class of matrices with **same margins, same order of groups**, the criterion $\preccurlyeq^D$ can be easily tested with via the orthant test:

Definition

The matrices $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ are **ordinal comparable** if (i) $\mathbf{1}_d^t \cdot \mathbf{A} = \mathbf{1}_d^t \cdot \mathbf{B}$ (same margins), (ii) all groups are ordered according to stochastic dominance in $\mathbf{A}$ and $\mathbf{B}$, and (iii) the order of the groups is the same in $\mathbf{A}$ and $\mathbf{B}$.

Implementation

▸ Back

Assessing $\mathbf{B} \preccurlyeq^D \mathbf{A}$ requires and infinity of Lorenz curve dominance comparisons.

For a special class of matrices with **same margins, same order of groups**, the criterion $\preccurlyeq^D$ can be easily tested with via the orthant test:

Definition

The matrices $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ are **ordinal comparable** if (i) $\mathbf{1}_d^t \cdot \mathbf{A} = \mathbf{1}_d^t \cdot \mathbf{B}$ (same margins), (ii) all groups are ordered according to stochastic dominance in $\mathbf{A}$ and $\mathbf{B}$, and (iii) the order of the groups is the same in $\mathbf{A}$ and $\mathbf{B}$.

Remark

Let $\mathbf{A} \in \mathcal{M}_d$ and $\mathbf{A}^*$ obtained from it through elimination of empty classes, split of classes, interchanges and permutation of groups operations, then $\mathbf{A}^* \sim^D \mathbf{A}$.

Implementation

Back

Assessing $\mathbf{B} \preccurlyeq^D \mathbf{A}$ requires and infinity of Lorenz curve dominance comparisons.

For a special class of matrices with **same margins, same order of groups**, the criterion $\preccurlyeq^D$ can be easily tested with via the orthant test:

Definition

The matrices $\mathbf{A},\ \mathbf{B} \in \mathcal{M}_d$ are **ordinal comparable** if (i) $\mathbf{1}_d^t \cdot \mathbf{A} = \mathbf{1}_d^t \cdot \mathbf{B}$ (same margins), (ii) all groups are ordered according to stochastic dominance in $\mathbf{A}$ and $\mathbf{B}$, and (iii) the order of the groups is the same in $\mathbf{A}$ and $\mathbf{B}$.

Remark

Let $\mathbf{A} \in \mathcal{M}_d$ and $\mathbf{A}^*$ obtained from it through elimination of empty classes, split of classes, interchanges and permutation of groups operations, then $\mathbf{A}^* \sim^D \mathbf{A}$.

Remark

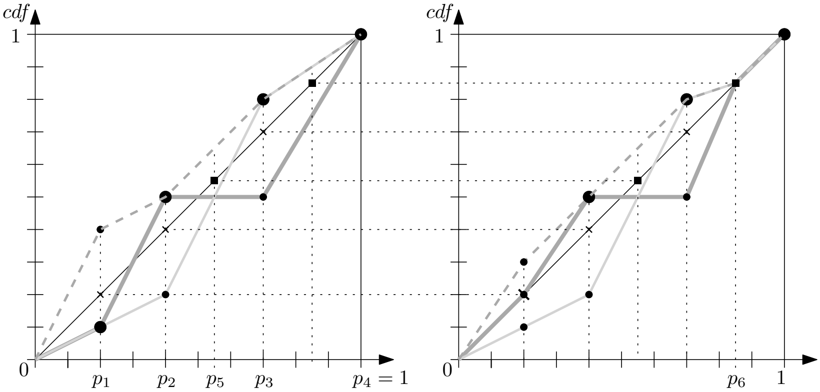
Any representation of the distributions $\overrightarrow{\mathbf{a}}(p)$ and $\overrightarrow{\mathbf{b}}(p)$ taken from matrices $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ can be equivalently obtained from two ordinal comparable matrices $\mathbf{A}^*, \mathbf{B}^* \in \mathcal{M}_d$ such that $\mathbf{A}^* \sim^D \mathbf{A}$ and $\mathbf{B}^* \sim^D \mathbf{B}$.

Implementation

Back

Example

$$\mathbf{A} = \begin{pmatrix} 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.4 & 0 & 0.5 \\ 0.1 & 0.1 & 0.6 & 0.2 \end{pmatrix} \quad \textit{and} \quad \mathbf{B} = \begin{pmatrix} 0.4{-}0.1 & 0.1{+}0.1 & 0.3 & 0.05 & 0.15 \\ 0.1{+}0.1 & 0.4{-}0.1 & 0 & 0.35 & 0.15 \\ 0.1 & 0.1 & 0.6 & 0.05 & 0.15 \end{pmatrix}$$

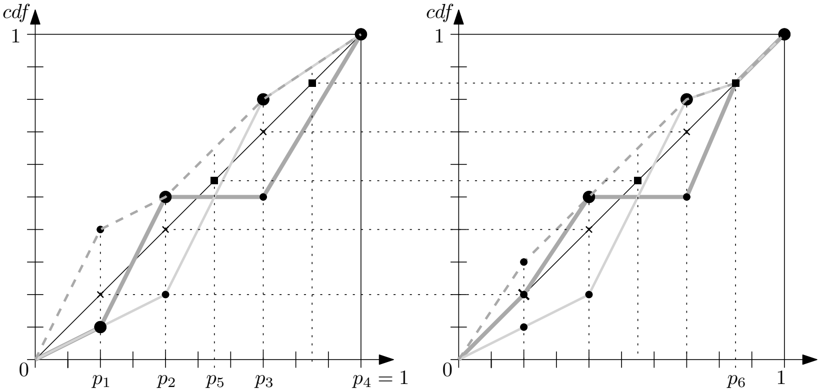


Implementation

Back

Example

$$\mathbf{A}^* = \begin{pmatrix} 0.4 & 0.1 & 0.15 & 0.15 & 0.1 & 0.1 \\ 0.1 & 0.4 & 0 & 0 & 0.25 & 0.25 \\ 0.1 & 0.1 & 0.3 & 0.3 & 0.1 & 0.1 \end{pmatrix} \textit{ and } \mathbf{B} = \begin{pmatrix} 0.3 & 0.2 & 0.15 & 0.15 & 0.0 \\ 0.2 & 0.3 & 0 & 0 & 0.3 \\ 0.1 & 0.1 & 0.3 & 0.3 & 0.0 \end{pmatrix}$$



Implementation

Back

Given that $\overrightarrow{b}_i(p) = \overrightarrow{b}_i^*(p)$ at any p, we have:

Corollary

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ *the following conditions are equivalent:*

(i) $\mathbf{B} \preccurlyeq^D \mathbf{A}$;

(ii) *There exist* $\mathbf{A}^*, \mathbf{B}^* \in \mathcal{M}_d$ *ordinal comparable that are obtained from* $\mathbf{A}$ *and* $\mathbf{B}$ *respectively through elimination of empty classes, split of classes, interchanges and permutation of groups operations, such that*

$$\Delta(h, p_j) := \sum_{i=1}^{h} \overrightarrow{b}^*_{(i)j} - \sum_{i=1}^{h} \overrightarrow{a}^*_{(i)j} \geq 0$$

for all $h = 1, \ldots, d$ *and for all* $j = 1, \ldots, n^*$.

Remark

When dissimilarity comparisons involve only two distributions (which we conventionally denote $h = 2$*), it can be shown that the test statistic* $\Delta(2, p_j)$ *coincides with*

$$\Delta(2, p_j) = \left|\overrightarrow{b^*}_{1j} - \overrightarrow{b^*}_{2j}\right| - \left|\overrightarrow{a^*}_{1j} - \overrightarrow{a^*}_{2j}\right| \leq 0, \quad \forall j.$$

Dissimilarity indices

Back

- Rank-dependent inequality indies (such as the Gini index family) can be used to value dissimilarity:

$$D_w(\mathbf{A}) := \int_0^1 \sum_{i=1}^{d} w_i(p)\overrightarrow{a}_{(i)}(p)dp$$

- Weighting functions $\mathcal{W}$:
  - $\sum_i w_i(p) = 0 \ \forall p$.
  - w_i non-decreasing in i
  - Example : S-Gini function

$$w_i(p) = \frac{1}{p}\left(1 - ((1 - \frac{i-1}{d})^k - (1 - \frac{i}{d})^k)\right)$$

for k a positive integer [Donaldson and Weymark 1980, Aaberge et al. 2019].

Corollary

For any $\mathbf{A}, \mathbf{B} \in \mathcal{M}_d$ the following conditions are equivalent:

(i) $\mathbf{B} \preccurlyeq^D \mathbf{A}$;

(ii) $D_w(\mathbf{B}) \ \leq \ D_w(\mathbf{A})$ *for all* $w \in \mathcal{W}$.

Empirical criterion

Back

Objective: derive empirical tests for dissimilarity comparisons involving distributions $F_i(y)$ for a group $i = 1, \ldots, d$ defined over cardinal outcomes $y \in \mathbb{R}$.

- Random sample of seize N, $\iota = 1, \ldots, N$ from the population
- N_i the sample size of group i
- y_ι an empirical occurrence for observation ι
- The index $j = 1, \ldots, n$ can be used to identify *distinct* empirical realizations $y(j)$
- Denote $a_{ij} = \sum_{\iota=1}^{N_i} \frac{1}{N_i} 1\{y_\iota = y(j)\}$. In small sample with no ties, likely $n = N$ and $a_{ij} \in \{1/N_i, 0\}$
- For any $y \in [y_j, y_{j+1}]$, denote the empirical cdf of group i:

$$\hat{F}_i(y) := \sum_{\iota=1}^{N_i} \frac{1}{N_i} 1\{y_\iota \leq y\} = \sum_{\iota=1}^{N_i} \frac{1}{N_i} 1\{y_\iota \leq y(j)\} = \overrightarrow{a}_{ij}.$$

- By linearity, $\hat{F}_i(y) \to^p F_i(y)$ and $\sum_i \frac{1}{d} \hat{F}_i(y) \to^p \overline{F}(y)$
- As the sample size N grows:

$$\overrightarrow{\mathbf{a}}(p) \ \longrightarrow^p \ \left(F_1(\overline{F}^{-1}(p)), \ldots, F_d(\overline{F}^{-1}(p)) \right), \ p \in [0, 1].$$

Dissimilarity preserving operations (ordered case)

Back

Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting* .

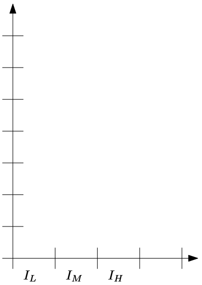
Dissimilarity preserving operations (ordered case)

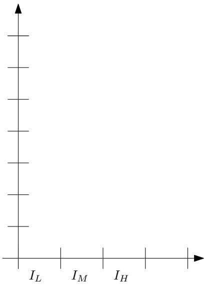
Back

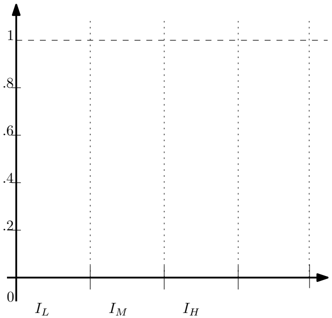
Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting .*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$







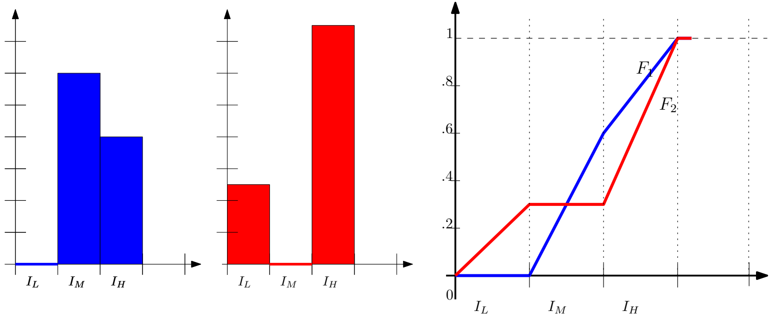
Dissimilarity preserving operations (ordered case)

Back

Axiom

Introduction of empty classes, split of classes, permutation of groups and Interchange preserve *the dissimilarity in the ordered setting* .

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$



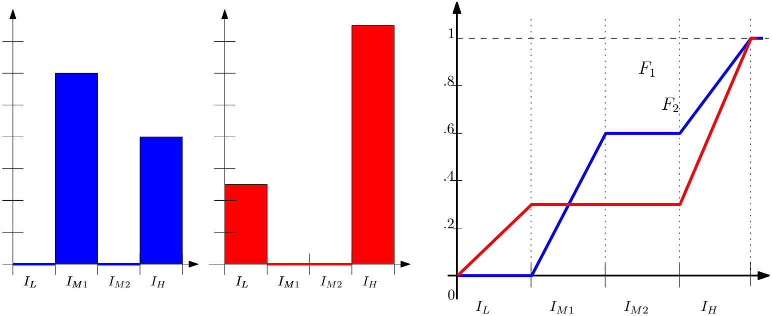
Dissimilarity preserving operations (ordered case)

Back

Axiom

Introduction of empty classes, split of classes, permutation of groups and Interchange preserve *the dissimilarity in the ordered setting* .

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$



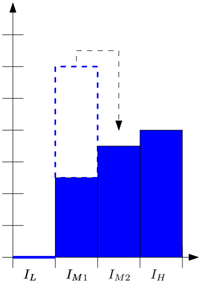
Dissimilarity preserving operations (ordered case)

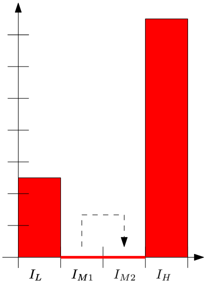
Back

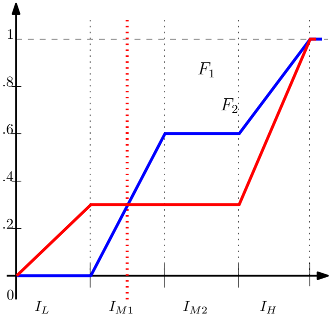
Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting* .

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$







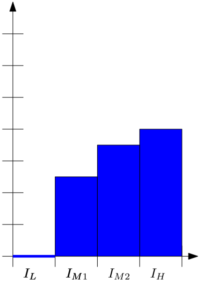
Dissimilarity preserving operations (ordered case)

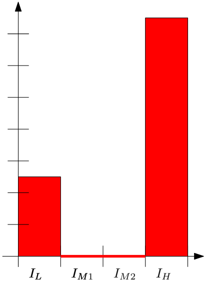
Back

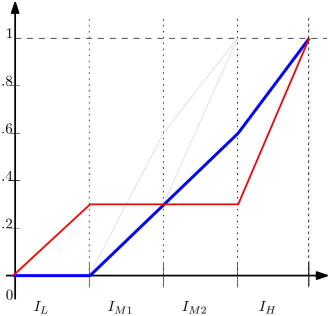
Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting .*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$







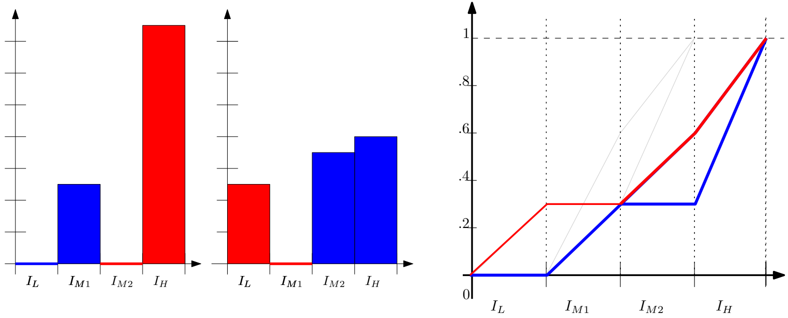
Dissimilarity preserving operations (ordered case)

Back

Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting* .

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.6 & 1 \\ 0.25 & 0.25 & 1 \\ X & X & X \end{pmatrix}$$



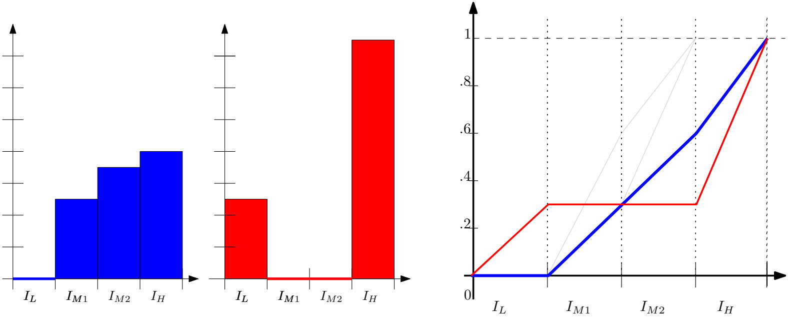
Dissimilarity preserving operations (ordered case)

Back

Axiom

Introduction of empty classes, **split of classes**, **permutation of groups and Interchange preserve** *the dissimilarity in the ordered setting* .

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$



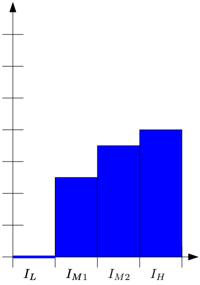
Dissimilarity preserving operations (ordered case)

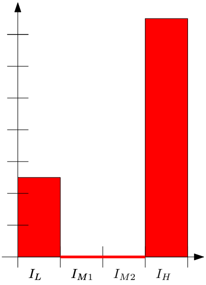
Back

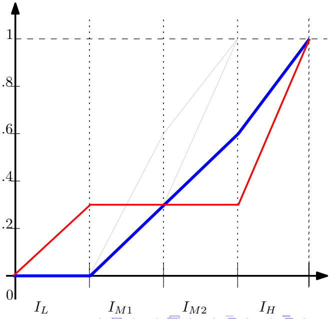
Axiom (**Exchange**)

*A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** and regressive for the **dominated group** reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$







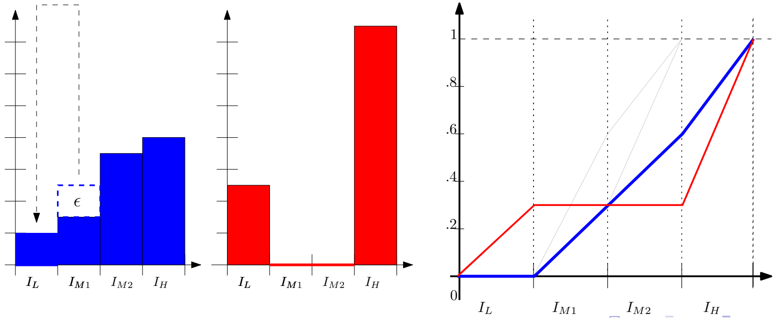
Dissimilarity preserving operations (ordered case)

Back

Axiom (**Exchange**)

A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$



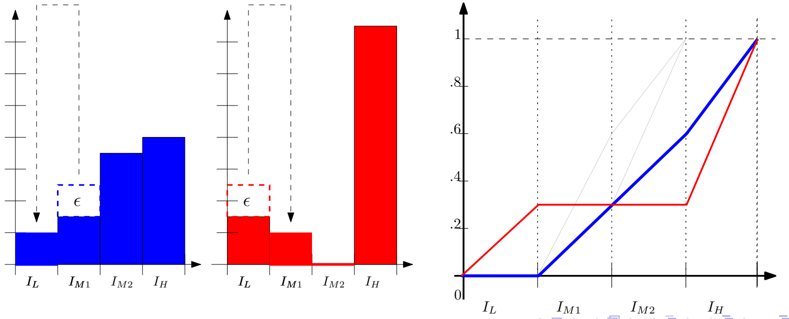
Dissimilarity preserving operations (ordered case)

Back

Axiom (**Exchange**)

A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$



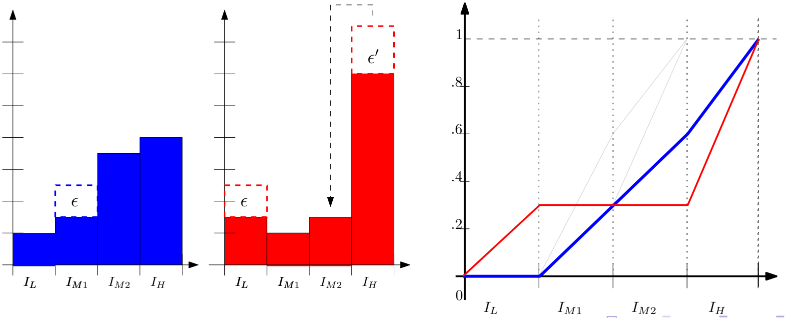
Dissimilarity preserving operations (ordered case)

Back

Axiom (**Exchange**)

A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$



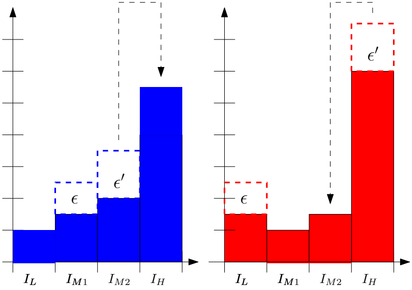
Dissimilarity preserving operations (ordered case)

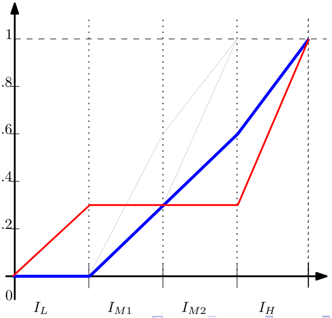
Back

Axiom (**Exchange**)

A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$





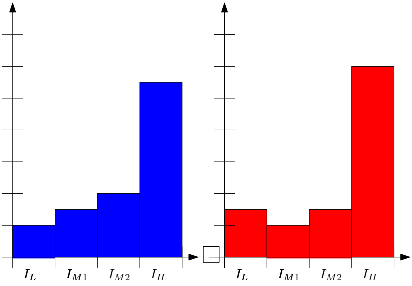
Dissimilarity preserving operations (ordered case)

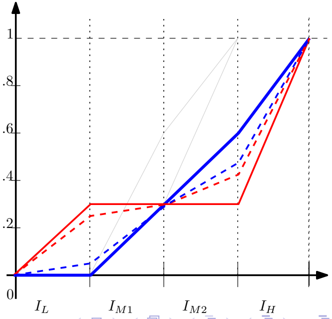
Back

Axiom (Exchange)

*A transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** and regressive for the **dominated group** reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0.35 & 0.4 \\ 0.25 & 0 & 0 & 0.75 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.6 & 1 \\ 0.25 & 0.25 & 0.25 & 1 \\ X & X & X & X \end{pmatrix}$$





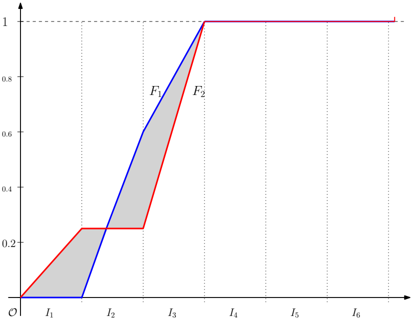
Geometry of dissimilarity

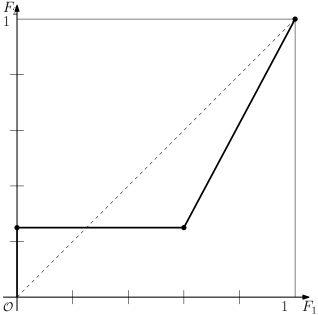
Back

- Consider a different mapping, the hypercube $[0,1]^d$
- The Monotone Path : $MP(\mathbf{A}) \;:=\; \left\{\overrightarrow{\mathbf{a}}(p) : p \in [0,1]\right\}$.
- Its expansion, the Path Polytope:

$$PP(\mathbf{A}) := \{\mathbf{z} \;:\; \mathbf{z} \in conv\left\{\mathbf{\Pi}_d \cdot \mathbf{p}\right\},\; \mathbf{\Pi}_d \in \mathcal{P}_d,\; \mathbf{p} \in MP(\mathbf{A})\}$$

.





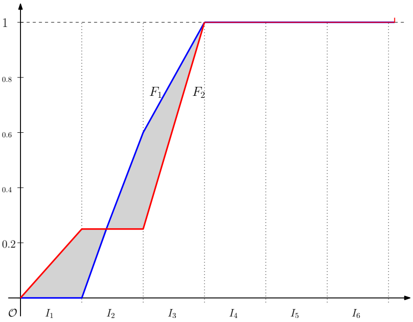
Geometry of dissimilarity

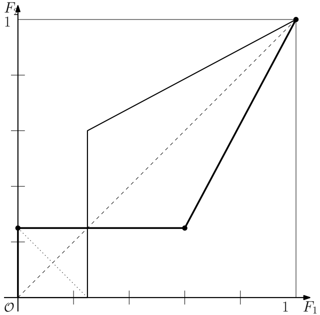
Back

- Consider a different mapping, the hypercube $[0,1]^d$
- The Monotone Path : $MP(\mathbf{A}) \ := \ \left\{ \overrightarrow{\mathbf{a}}(p) : p \in [0,1] \right\}$.
- Its expansion, the Path Polytope :

$$PP(\mathbf{A}) := \{\mathbf{z} \ : \ \mathbf{z} \in conv\left\{\mathbf{\Pi}_d \cdot \mathbf{p}\right\}, \ \mathbf{\Pi}_d \in \mathcal{P}_d, \ \mathbf{p} \in MP(\mathbf{A})\}$$

.





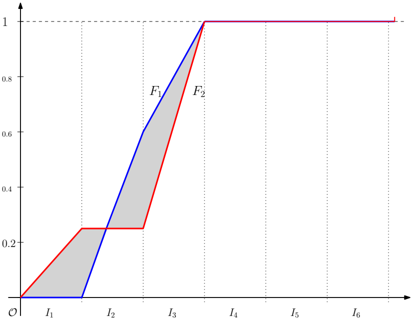
Geometry of dissimilarity

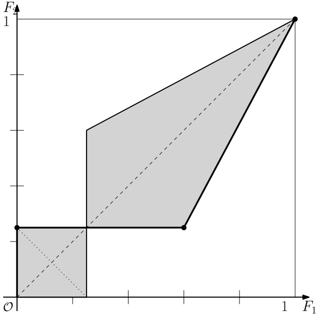
Back

- Consider a different mapping, the hypercube $[0,1]^d$
- The Monotone Path : $MP(\mathbf{A}) \ := \ \left\{ \overrightarrow{\mathbf{a}}(p) : p \in [0,1] \right\}$.
- Its expansion, the Path Polytope :

$$PP(\mathbf{A}) := \{ \mathbf{z} \ : \ \mathbf{z} \in conv \{ \mathbf{\Pi}_d \cdot \mathbf{p} \}, \ \mathbf{\Pi}_d \in \mathcal{P}_d, \ \mathbf{p} \in MP(\mathbf{A}) \}$$

.

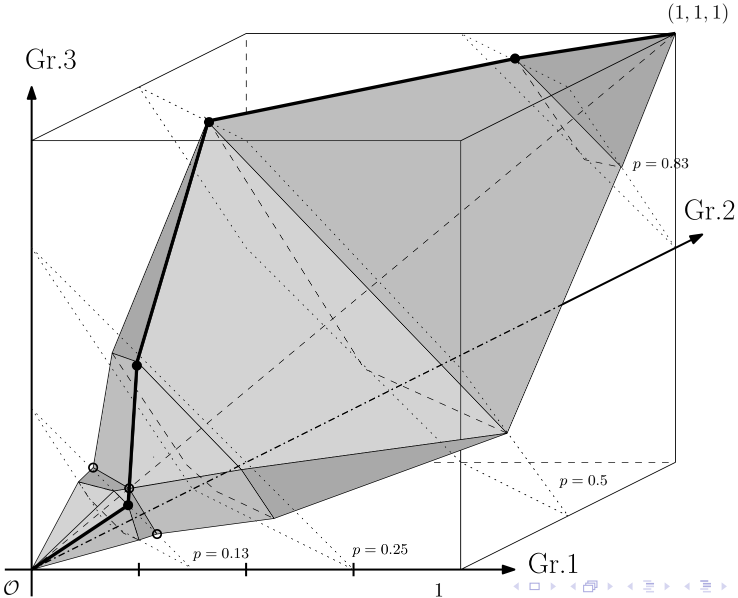




Geometry of dissimilarity

▸ Back

- Complex to draw for $d > 2$, easy to test.



Geometry of dissimilarity

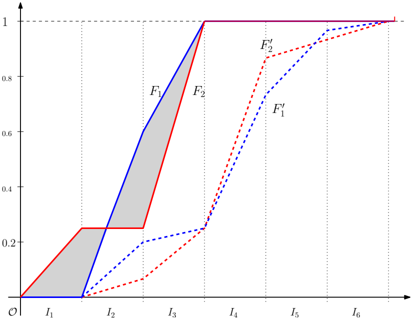
Back

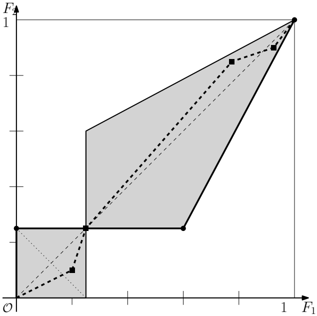
- The dissimilarity test: $PP(\mathbf{B}) \subseteq PP(\mathbf{A})$.

Geometry of dissimilarity

Back

- The dissimilarity test: $PP(\mathbf{B}) \subseteq PP(\mathbf{A})$.
  - $\subseteq$-ordering induce a partial ranking
  - diagonal $= PP(\mathbf{S}) \subseteq PP(\mathbf{A}) \subseteq PP(\mathbf{D}) =$ hypercube

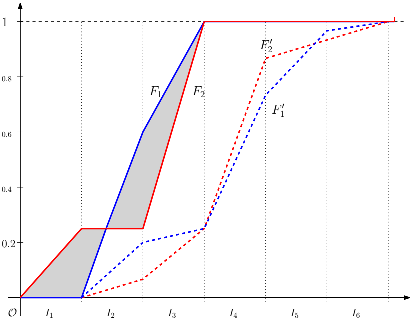


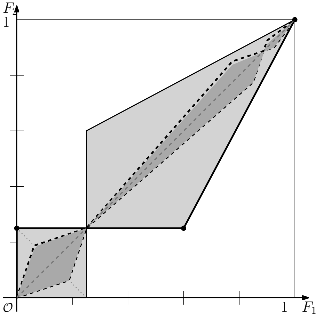


Geometry of dissimilarity

Back

- The dissimilarity test: $PP(\mathbf{B}) \subseteq PP(\mathbf{A})$.
  - $\subseteq$-ordering induce a partial ranking
  - diagonal $= PP(\mathbf{S}) \subseteq PP(\mathbf{A}) \subseteq PP(\mathbf{D}) =$ hypercube

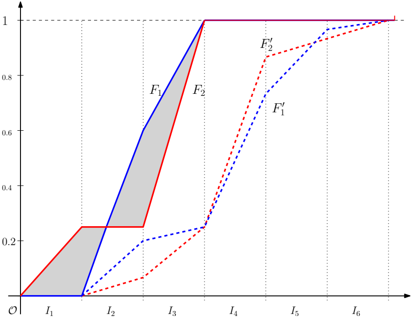


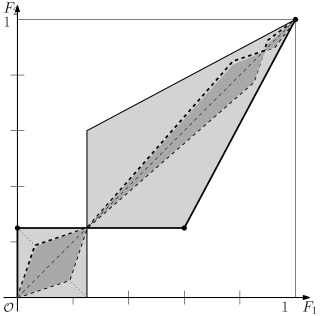


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

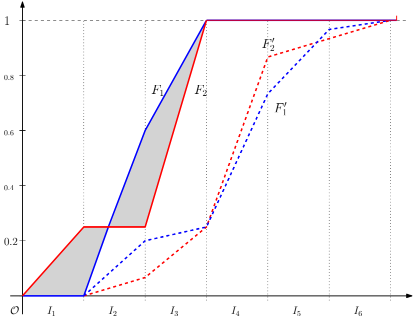


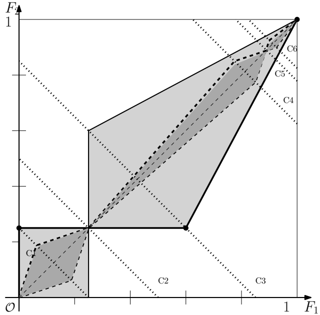


Implementation

▸ Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

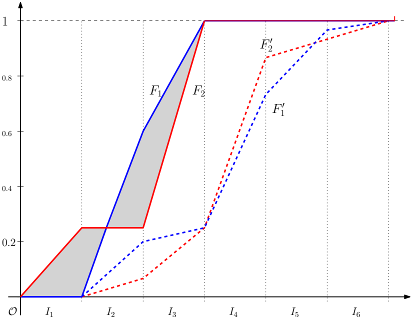


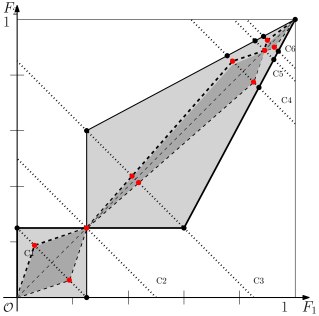


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

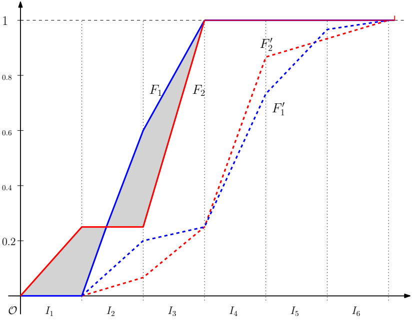


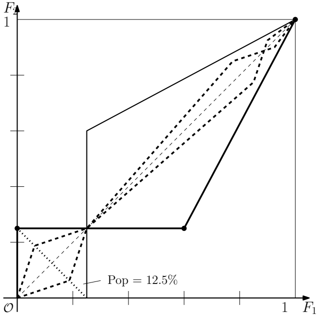


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

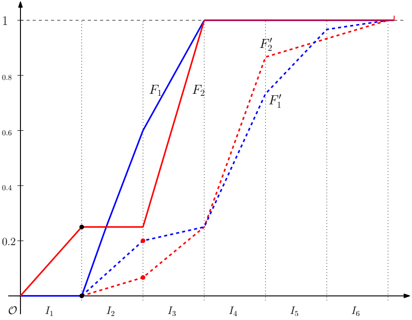


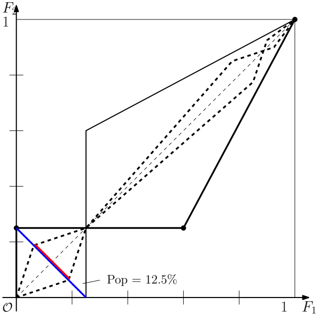


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

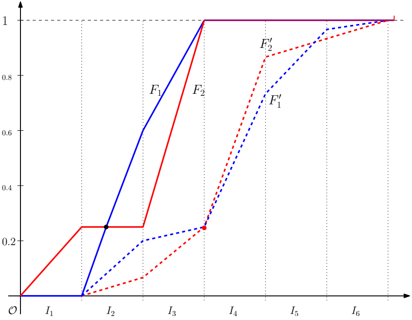


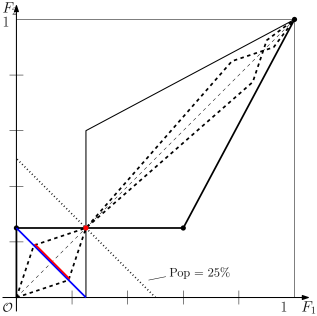


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

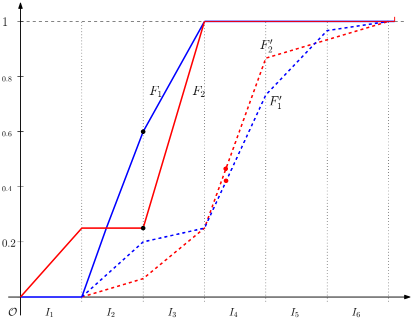


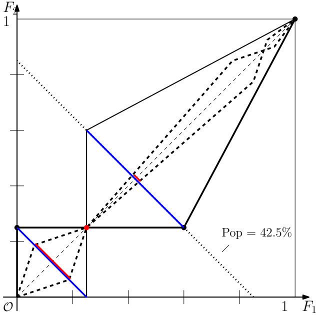


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

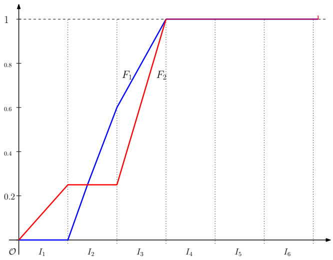


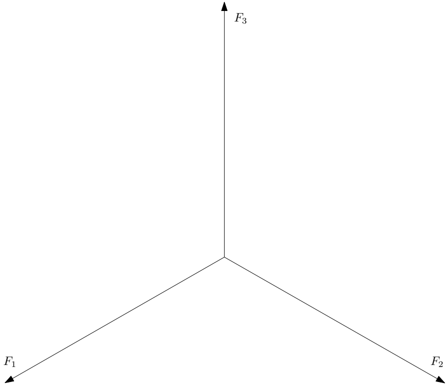


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

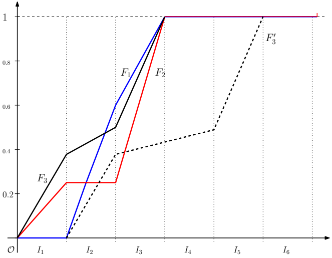


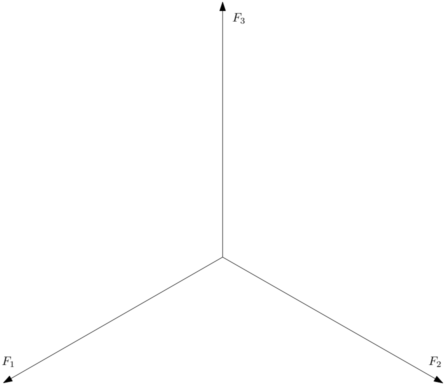


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

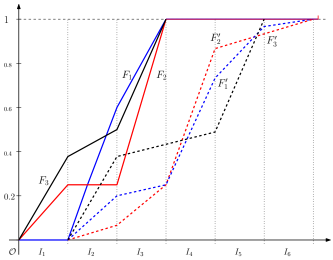


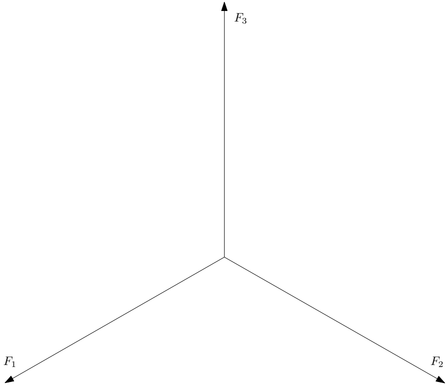


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

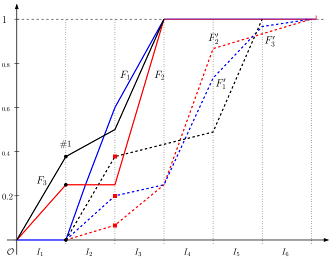


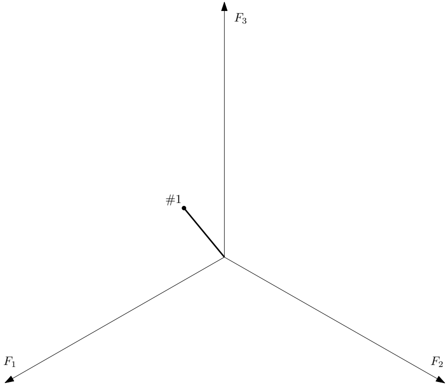


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

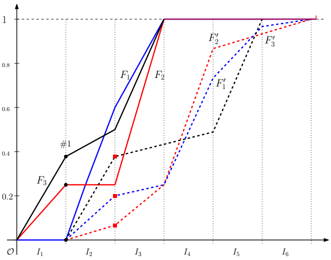


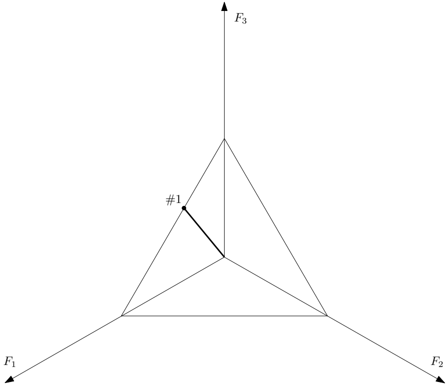


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

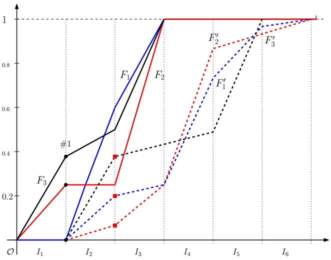


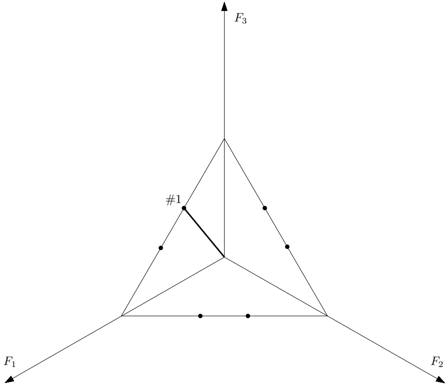


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

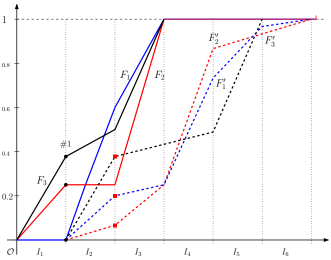


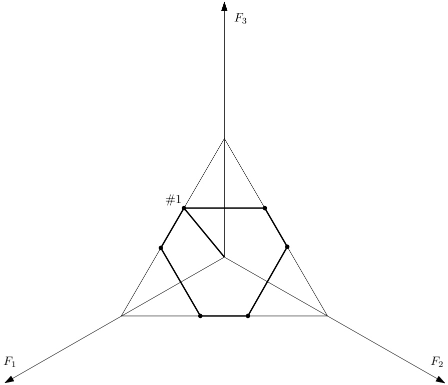


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

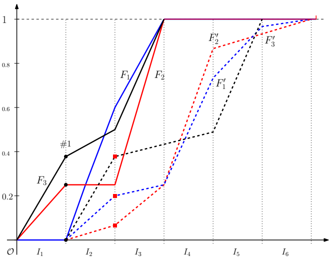


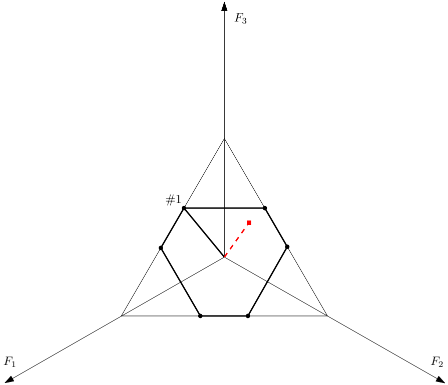


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

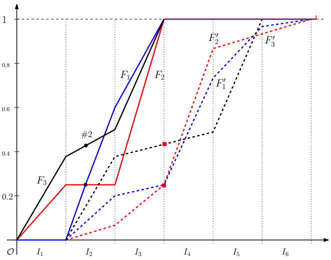


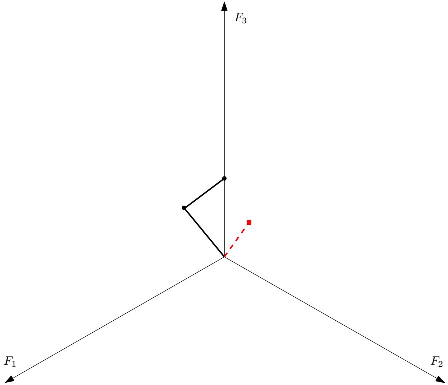


Implementation

▸ Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

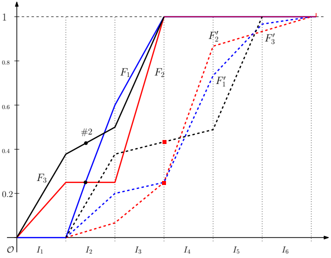


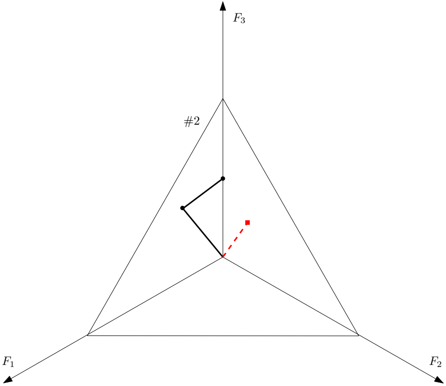


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

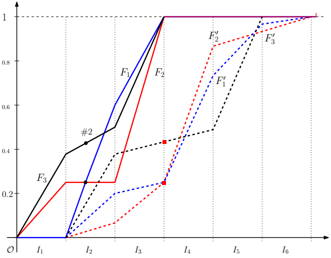


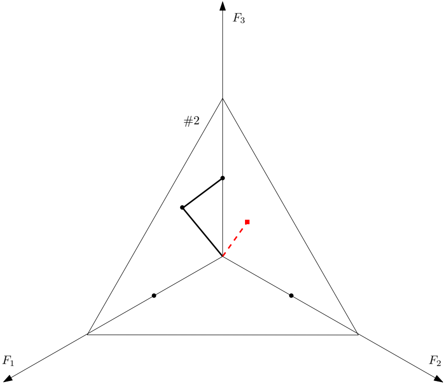


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

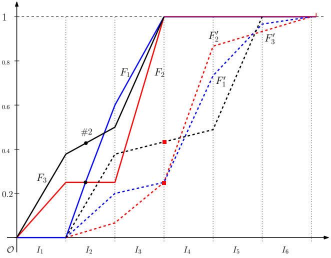


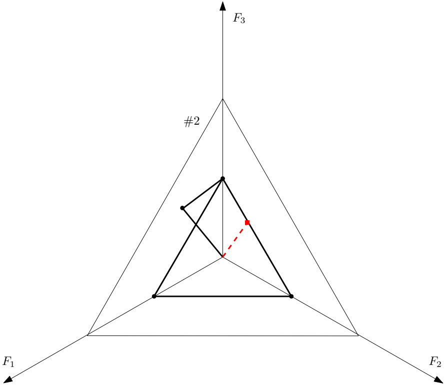


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

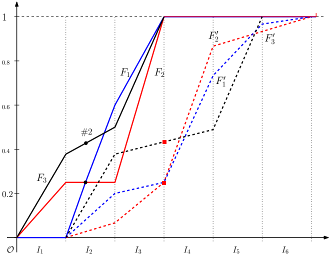


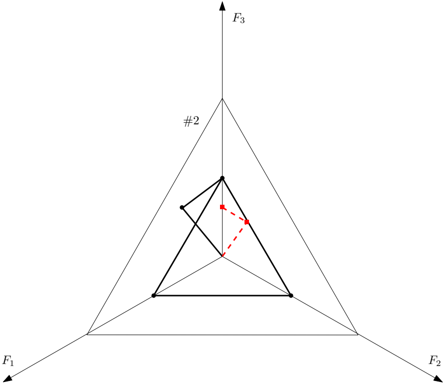


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

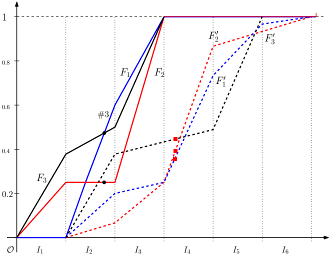


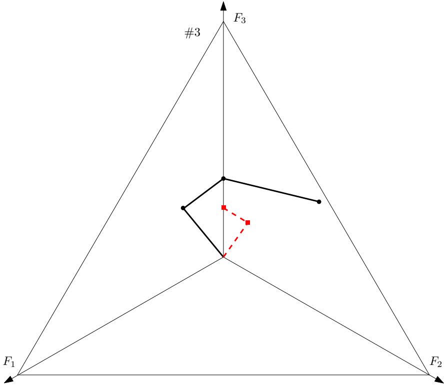


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.

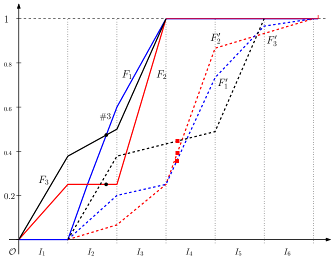


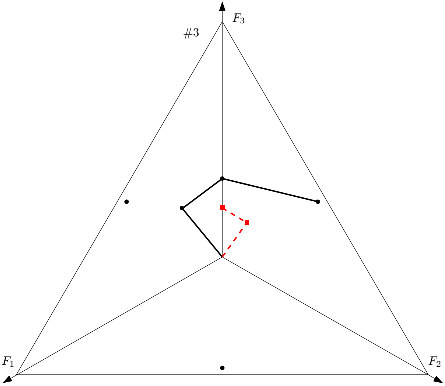


Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.





Implementation

Back

- Main idea: it is sufficient to test Lorenz dominance on selected intercepts.
  - $d = 2$ is visual.
  - $d > 2$ needs some effort... and the **Kolm's (1969) triangles**.