Ranking distributions of an ordinal attribute

Nicolas Gravel, Aix-Marseille Univ. (Aix-Marseille School of economics), CNRS & EHESS

Brice Magdalou, Université de Montpellier, LAMETA

Patrick Moyes, CNRS, Gretha, University of Bordeaux IV

Canazei winter school, January 2015

Object of the talk :

- Foundations of normative evaluation.
- Classical setting: comparing distributions of a **cardinally measurable attribute** (income) between a given number (n say) of households.
- **Q**: When can we say that distribution $x = (x_1, \ldots, x_n)$ is «unquestionably better » than $y = (y_1, \ldots, y_n)$?
- The classical literature (Kolm 1966, Atkinson 1970, Dasgupta Sen & Starett (1973), Sen (1973), etc. has provided **3** equivalent answers to that question.

3 equivalent answers (Hardy, Littlewood Polya (HLP))

- **1)** When all utilitarian social planners who assume that individuals convert income into utility by the same increasing and concave function would agree so.
- 2) When x has been obtained from y by a finite sequence of permutations and/or Pigou-Dalton transfers and/or increments.
- **3)** When the generalized Lorenz curve associated to x is never below that associated to y.

This equivalence is nice because it connects together:

- An **explicit** and **robust ethical foundation** (utilitarian (actually even larger, see Gravel & Moyes 2013) unanimity over a plausible class of individual utility functions.
- **Elementary transformations** (Pigou-Dalton transfers, increments and permutations) that identify clearly the nature of the normative improvements at stake.
- **Empirically implementable** criteria (Lorenz dominance) that can be (and are!) used in practice to perform normative evaluation.
- **This equivalence is foundational:** When more specific (and controversial) inequality or aggregate indices are used, their consistency with any of these equivalent partial answer is considered fundamental.

This research:

- Is concerned with establishing analogous foundations to the problem of comparing distributions of an interpersonally comparable **ordinal** attribute between households.
- Ordinal attribute: attribute whose **meaningful** numerical measurement is unique up to an increasing transformation.
- $(x_1,\ldots,x_n)$ is better than $(y_1,\ldots,y_n)$ if and only if $(f(x_1),\ldots,f(x_n))$ is better than $(f(y_1),\ldots f(y_n))$ for any increasing function f.
- Examples: access to housing, self-reported happiness (??), health, body mass index, Pisa scores, years of schooling, IQ, bibliometric indices, etc.

Ordinal attribute

- When can we say that a distribution $x = (x_1, \ldots, x_n)$ of an ordinal attribute is «unquestionably better » than distribution $y = (y_1, \ldots, y_n)$?
- Many researchers answer this question in just the same way as for a cardinal attribute.
- **Problem**: What is the meaning of a Pigou-Dalton transfer of an ordinal attribute between two individuals ?
- Equivalently: what is the meaning of adding quantities of the ordinal attribute in the way required by the construction of Lorenz curves ?
- Notice carefully: no problem with increments and permutations! The difficulty comes with Pigou-Dalton, and the meaning of "inequality reduction" with an ordinal variable.

Ordinal attribute:

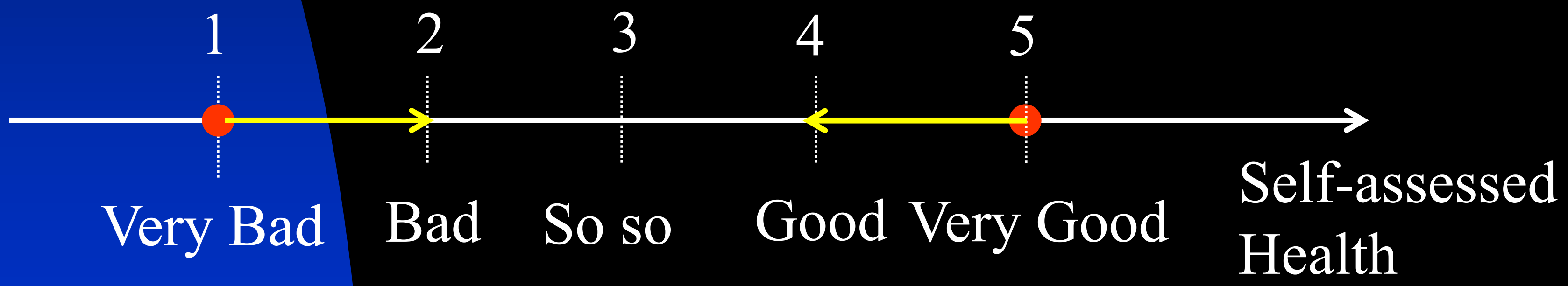
- Some important exceptions in the literature explicitly recognize the specificities of ordinal measurement.
- Alison & Foster (AF) (J. Health. E, 2004)
- AF: applies to distributions with the same median, and assumes that the attribute can only take on finitely many values.
- AF: x is better than y if the cumulative distribution of the population in x lies everywhere below that of y below the median and everywhere above that of y above the median.
- Abul-Naga & Yalcin (JHealthE, 2008): develop and apply indices that are compatible with AF.
- Cowell & Flachaire (2014): develop and implement indices that are expressed in terms of distance from an exogenous reference ideal (perhaps the median)

In this research:

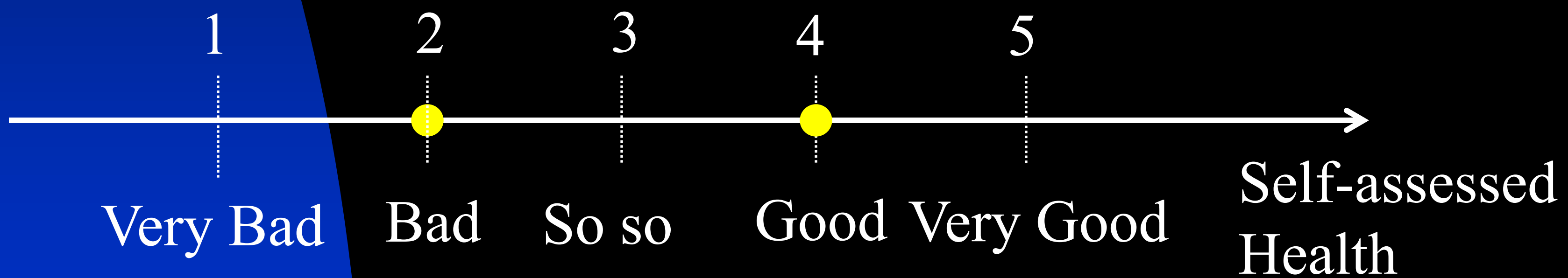
- We consider a transfer principle – due to Peter J. Hammond (Econometrica 1976) - that is arguably more appropriate than Pigou-Dalton for capturing our intuition about meaning of inequality reduction in an ordinal setting.
- We establish an analogue to the HLP theorem for that transfer principle (along with increments and permutations).
- Specifically, we establish the equivalence between the Hammond transfer principle and
- (*i*) a normative dominance criterion (represented by an additively separable objective function)
- (*ii*) An easy-to-use implementable dominance criterion.
- As in AF, our analysis is limited to an ordinal attribute that can take finitely many different values (categories)

Hammond transfer

Hammond transfer

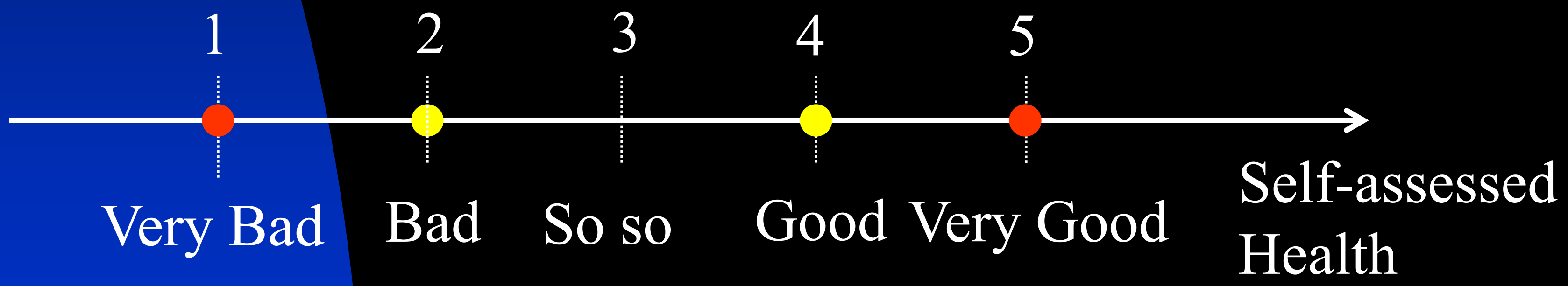


Hammond transfer

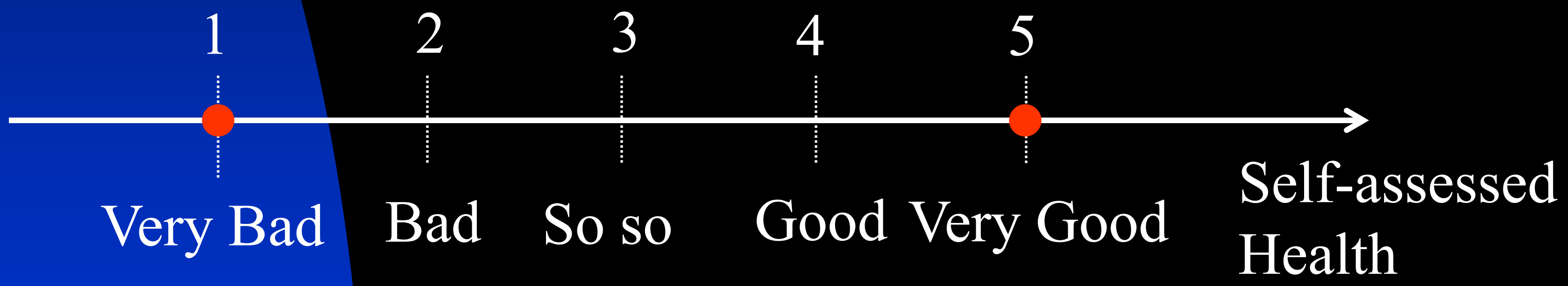


Hammond transfer

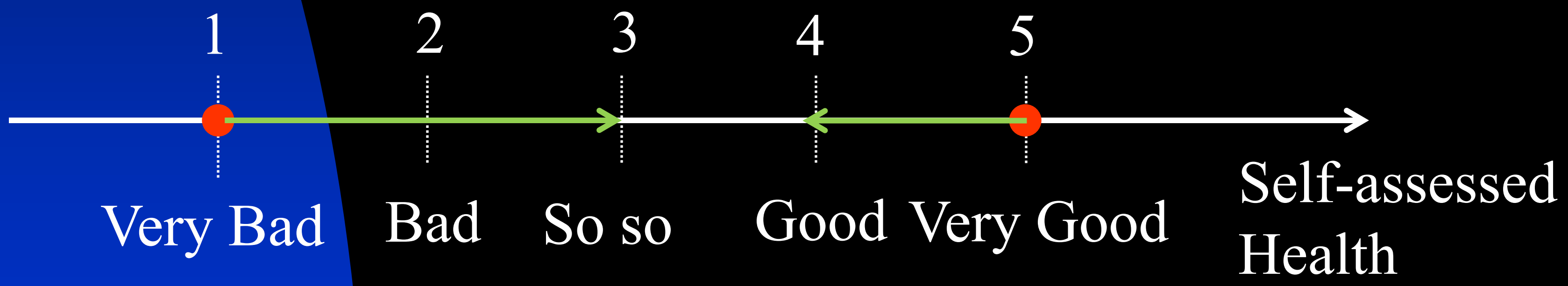
The yellow distribution has been obtained from the red one by a Hammond transfer (that is also a Pigou-Dalton Transfer).



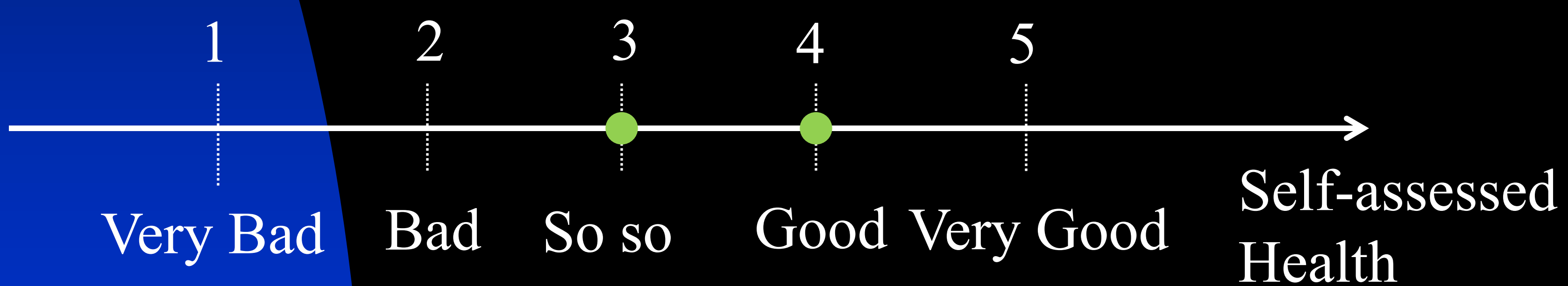
Hammond transfer



Hammond transfer

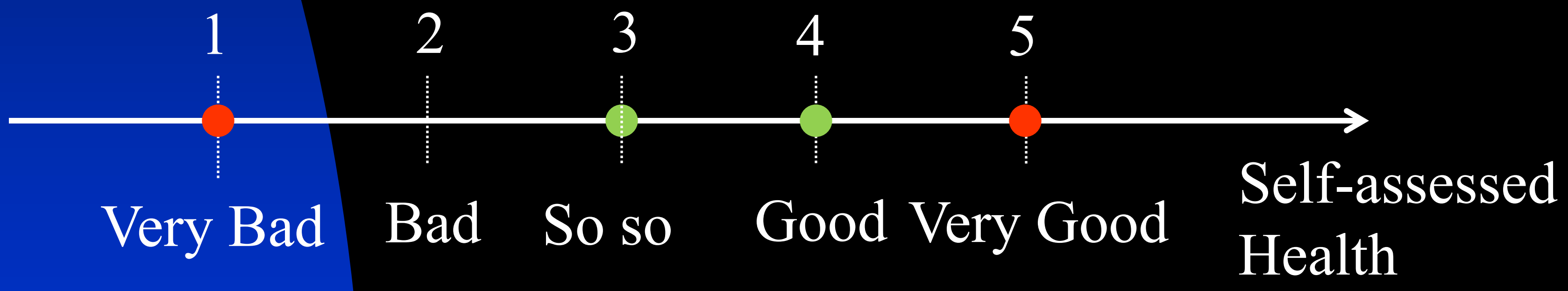


Hammond transfer

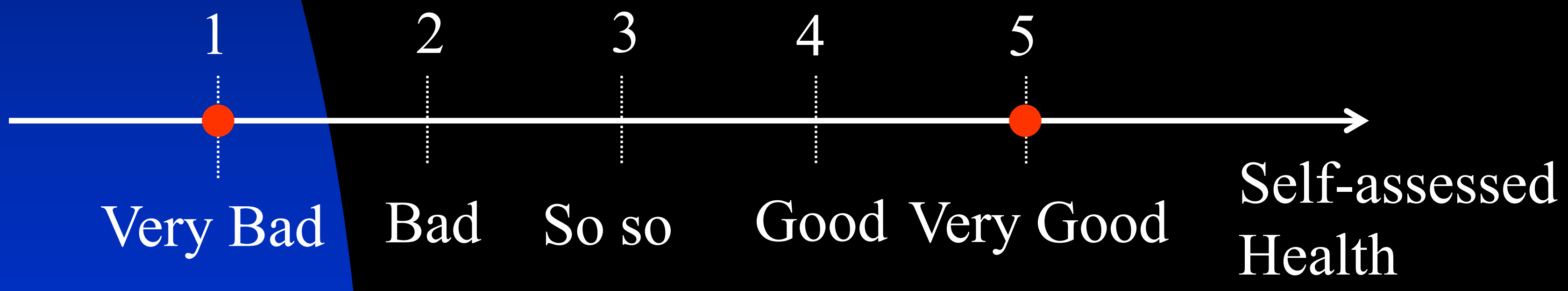


Hammond transfer

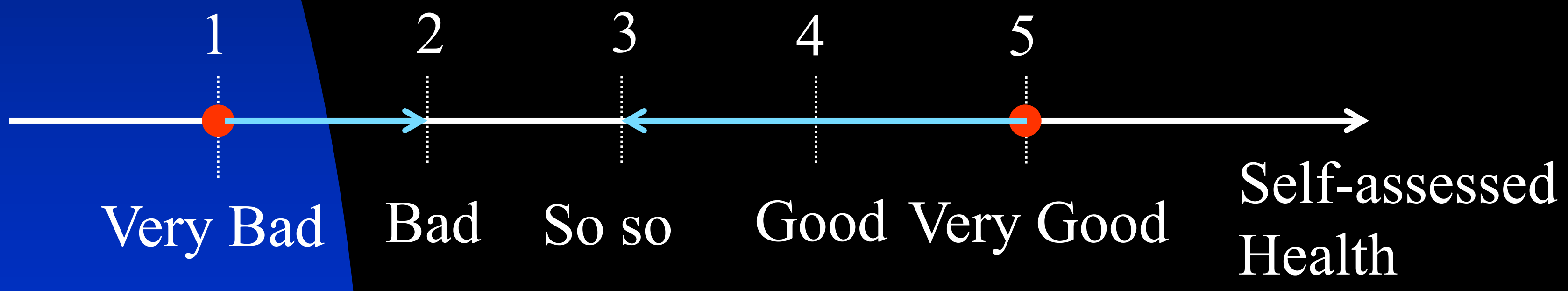
The green distribution has been obtained from the red one by a Hammond transfer (that is not a Pigou-Dalton transfer).



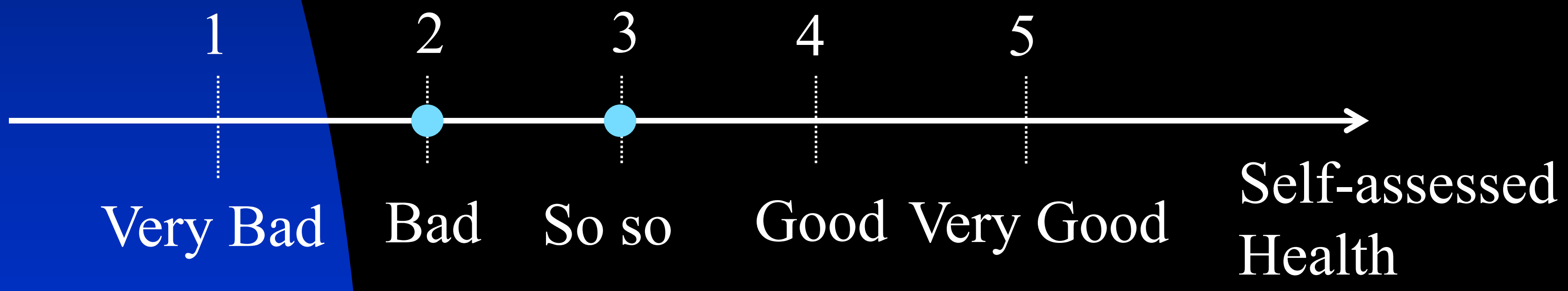
Hammond transfer



Hammond transfer

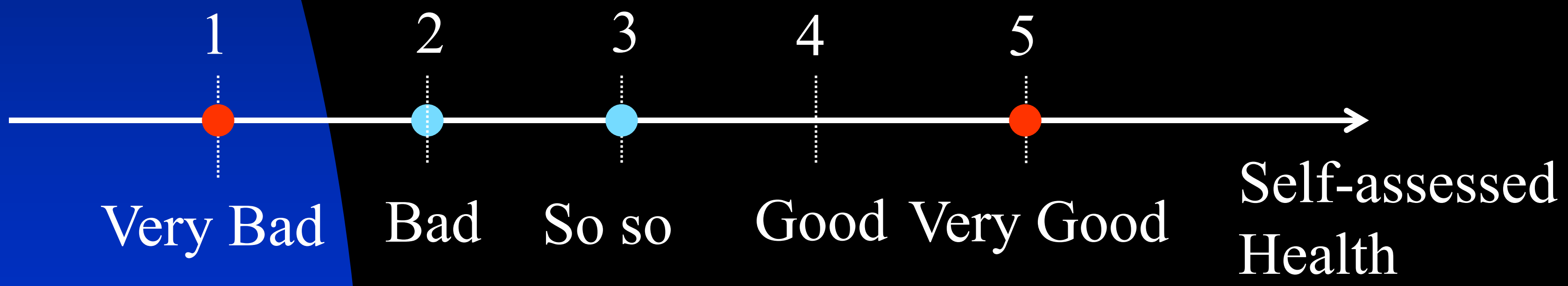


Hammond transfer



Hammond transfer

The blue distribution has been obtained from the red one by a Hammond transfer (that is not a Pigou-Dalton transfer).



Normative dominance:

- Consider an ordinal attribute that can take k different values (categories) ordered from the worst to the best.
- Assume there are n individuals.
- Under an anonymity principle (individual names don't matter), a distribution (d say) of the ordinal attribute can be described by the k-tuple $(n_1^d,\dots,n_k^d)$ where n_i^d denote the number of individuals in d who are in category i.
- **Normative criterion**: d is better than d' if and only if:

$$\sum_{j=1}^{k} n_j^d \alpha_j \geq \sum_{j=1}^{k} n_j^{d'} \alpha_j$$

for some numbers $\alpha_1,\dots,\alpha_k$ representing the (utility ?) values attached to the categories by the « planner »
(presumably $\alpha_1 \leq \dots \leq \alpha_k$)

Why normatively evaluating distributions of an ordinal attribute in this fashion ?

- Utilitarianism (in fact average utilitarianism if one assumes a variable population).
- Non-welfarist justification (see Gravel, Marchand, Sen (2011))
- **Question**: would such a criterion be sensitive to Hammond transfers ?
- **Answer:** yes if (and only if) the (utility) numbers $\alpha_1,\ldots,\alpha_k$ satisfy the following property (*p2*):

$\alpha_h - \alpha_g \geq \alpha_j - \alpha_i$ for all $g \leq h \leq i \leq j$

(strong concavity ?)

Normative dominance

- d is better than d' if and only if:

$$\sum_{j=1}^{k} n_j^{d} \alpha_j \geq \sum_{j=1}^{k} n_j^{d'} \alpha_j$$

For all numbers $\alpha_1,\ldots,\alpha_k$ satisfying $\alpha_1 \leq \ldots \leq \alpha_k$ and *p2*

Implementable tool: H^+-curve

- Consider a distribution d
- For every i=1,…,k, denote by $F(i,d)$ the number of people in d who have i or less of the attribute:

$$F(d,i)=\sum_{h=1}^{i} n_h^d$$

CDF in statistics (up to the division by the population size)

Our H^+-curve is constructed recursively from the F curve as follows:

$$H^+(d,1)=F(d,1)$$

and

$$H^+(d,i)=2H^+(d,i-1)+n_i^d$$

for i=2,…,k

The H^+-curve

- To the very best of our knowledge, it is a new statistical tool.
- Easy to apply and work like the CDF: start with the number (fraction) of people in the lower category, double that number (fraction) and add the number (fraction) of people in the immediately superior category, double the number of the preceding step and add the number of people in the next immediately superior category and so on…
- Two alternative ways to define the H^+-curve (for $i = 2,\ldots,k$):

$$H^+(d,i) = \sum_{h=1}^{i} (2^{i-h}) n_h^d$$

and

$$H^+(d,i) = \sum_{h=1}^{i-1} (2^{i-h-1}) F(d,h) + F(d,i)$$

The H^+-curve

- As established in this work, the comparisons of H^+-curves is very closely related to the possibility of going from one distribution to the other by a finite sequence of permutations and/or increment and/or Hammond transfers.
- $H^+(d,i) \leq H^+(d',i)$ for all i (with at least one strict inequality) if and only if one can go from d' to d by a finite sequence of permutations and/or increments and/or Hammond transfers.

Formal definition of an increment

- Distribution d is obtained from d' by an increment if there exists a category $j < k$ such that:

For all l distinct from j and

$$n_j^d = n_j^{d'} - 1; n_{j+1}^d = n_{j+1}^{d'} + 1$$

Formal definition of Hammond Transfer

- Distribution d is obtained from d' by a Hammond transfer if there are four categories $1 \leq g < h \leq i < j$ such that:

For all l distinct from g, h, i and j

$$n_g^d = n_g^{d'} - 1; n_h^d = n_h^{d'} + 1$$

and

$$n_i^d = n_i^{d'} + 1; n_j^d = n_j^{d'} - 1$$

Main theorem:

- The following three statements are equivalent:
- 1: $H^{+}(d,i) \leq H^{+}(d',i)$ for all i
- 2: d has been obtained from d' by a finite sequence of increments and/or Hammond transfers.
- 3: $\sum_{j=1}^{k} n_j^{d} \alpha_j \geq \sum_{j=1}^{k} n_j^{d'} \alpha_j$ For all numbers $\alpha_1,\ldots,\alpha_k$ satisfying $\alpha_1 \leq \ldots \leq \alpha_k$ and *p2*

Remarks on this result

- It does not disentangle increments on the one-hand and Hammond-transfers and permutations on the other.
- Increments deal with « efficiency » while Hammond transfers deal with « equity ».
- Question: Could we design an operational criterion that identifies only equity ?

Identifying pure ordinal equalization

- How do we proceed in the standard cardinal case ?
- Distribution x has been obtained from distribution y by a finite sequence of Pigou-Dalton transfers if and only if the sum of the i lowest income is larger in x than in y whatever i is ***and*** x and y have the same mean.
- Equivalently (Marshall & Olkin 1979): x has been obtained from distribution y by a finite sequence of Pigou-Dalton transfers if and only if, for every i , the sum of the i lowest incomes is larger in x than in y ***and*** the sum of the i highest incomes is lower in x than in y

Identifying pure ordinal equalization

- The first definition is of no clear interest in an ordinal setting where the notion of two distributions having the same mean is vacuous.
- The second one (based on the intersection of two dual quasi-orderings) is perhaps more promising.
- Indeed, we do have a characterization of a « dual » H- dominance criterion that is of the same spirit than the cardinal test of cumulating incomes from the top (rather than from the bottom) and of trying to get these cumulate income smaller (rather than bigger).

A dual curve: the H^--curve

- Consider a distribution d
- For every i=1,…,k-1, denote by $S(i,d)$ the number of people in d who have i or more of the attribute:

$$S(d,i) = n - \sum_{h=1}^{i} n_h^d = \sum_{h=i+1}^{k} n_h^d$$

CCDF in statistics (called also the « survival » function)

H^-curve: constructed recursively from the S curve as follows:

$$H^-(d,k-1) = S(d,k-1)$$ and

$$H^-(d,i) = 2H^-(d,i+1) + n_{i+1}^d$$ for i=1,…,k-2

The H$^-$-curve

- Similar recursion than for the $H+$ curve (but starting from above): start with the number (fraction) of people in the highest category, double that number (fraction) and add the number (fraction) of people in the immediately inferior category, double the number of the preceding step and add the number of people in the next immediately inferior category and so on…
- Two alternative ways to define the H^--curve (for $i = 1,\ldots,k-1$):

$$H^-(d,i) = \sum_{h=i+1}^{k} (2^{h-i-1}) n_h^d$$

and (for $i = 1,\ldots,k-2$):

$$H^-(d,i) = \sum_{h=i+1}^{k-1} (2^{h-i-1}) S(d,h) + S(d,i)$$

Dual theorem:

- The following three statements are equivalent:
- 1: $H^{-}(d,i) \leq H^{-}(d',i)$ for all $i=1,\ldots,k-1$
- 2: d has been obtained from d' by a finite sequence of decrements and/or Hammond transfers and/or permutations.
- 3: $$\sum_{j=1}^{k} n_j^{d} \alpha_j \geq \sum_{j=1}^{k} n_j^{d'} \alpha_j$$ For all numbers $\alpha_1,\ldots,\alpha_k$ satisfying $\alpha_k \leq \ldots \leq \alpha_1$ and $p2$

Ordinal pure inequality reduction ?

- The intersection of H^- and H^+ dominance ?
- The transitive closure of the ranking « being obtained by a Hammond transfer » belongs to that intersection.
- What else ?
- We conjecture that the answer to this question is: nothing!
- We have the following result that makes us optimistic (but not sure yet)about the truth of this conjecture.

Ordinal pure inequality reduction ?

- Theorem: Consider the following three statements:
- 1) $H^{-}(d,i) \leq H^{-}(d',i)$ and $H^{+}(d,i) \leq H^{+}(d',i)$ for all $i=1,\ldots,k-1$
- 2) $\sum_{j=1}^{k} n_j^{d}\alpha_j \geq \sum_{j=1}^{k} n_j^{d'}\alpha_j$ For all numbers $\alpha_1,\ldots,\alpha_k$ satisfying *p2*
- 3) d has been obtained from d' by a finite sequence of Hammond transfers and/or permutations.
- Statements 1) and 2) are equivalent and are both implied by 3).
- We don't know yet wether the reverse implication holds

Empirical illustrations

- These H dominance notions are very easy to use.
- Let us illustrate this:
- first with data on health inequality in Switzerland (taken from Abul-Naga and Yalcin, J. of Health E., 2008)
- Second with French data on the distribution of body mass index

Distribution of self-reported health status in Switzerland (5 categories)

Region	n_1/n	n_2/n	n_3/n	n_4/n	n_5/n
Léman	0.01	0.04	0.11	0.56	0.28
North-West	0.01	0.04	0.13	0.63	0.19
Central	0.00	0.02	0.11	0.63	0.24
Middle-Land	0.01	0.03	0.13	0.60	0.23
East	0.00	0.03	0.11	0.64	0.22
Ticino	0.01	0.05	0.11	0.70	0.13
Zurich	0	0.03	0.10	0.65	0.22

Distribution of self-reported health status in Switzerland (5 categories)

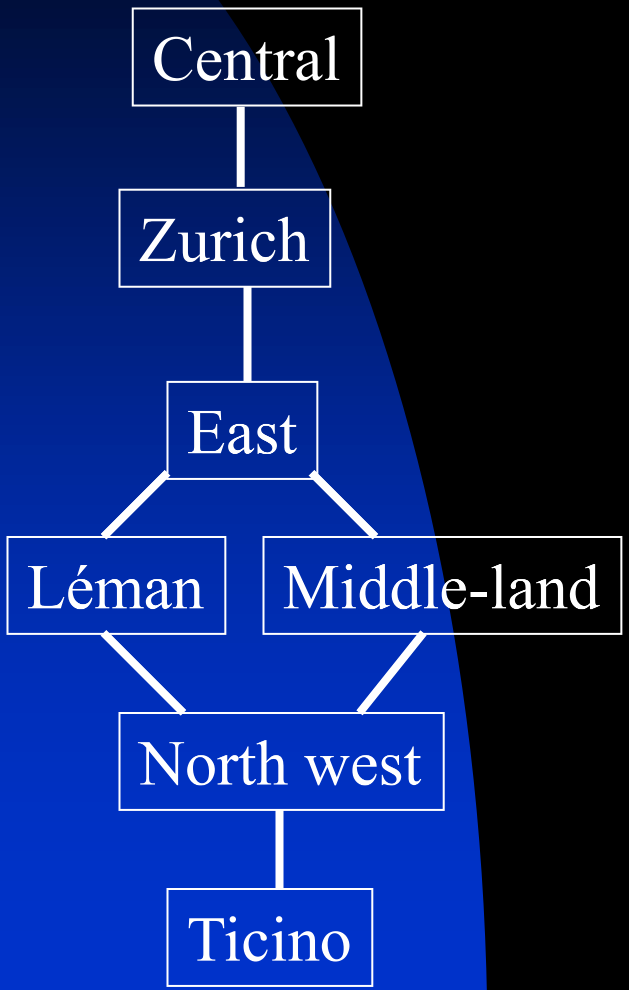
Region	$F(d,1)$	$F(d,2)$	$F(d,3)$	$F(d,4)$	$F(d,5)$
Léman	0.01	0.05	0.16	0.72	1.00
North-West	0.01	0.05	0.18	0.81	1.00
Central	0.00	0.02	0.13	0.76	1.00
Middle-Land	0.01	0.04	0.17	0.77	1.00
East	0.00	0.03	0.14	0.78	1.00
Ticino	0.01	0.06	0.17	0.87	1.00
Zurich	0	0.03	0.13	0.78	1.00

Distribution of self-reported health status in Switzerland (5 categories)

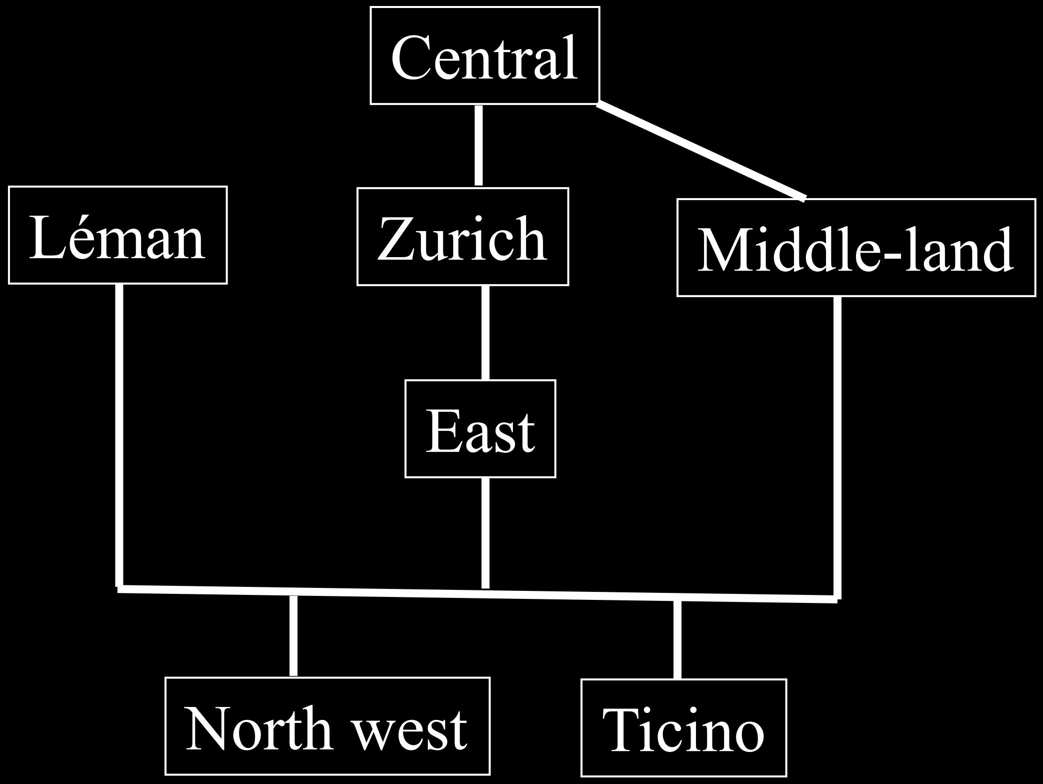
Region	$H(d,1)$	$H(d,2)$	$H(d,3)$	$H(d,4)$	$H(d,5)$
Léman	0.01	0.06	0.23	1.02	2.32
North-West	0.01	0.06	0.25	1.13	2.45
Central	0.00	0.02	0.15	0.93	2.10
Middle-Land	0.01	0.05	0.23	1.06	2.35
East	0.00	0.03	0.17	0.98	2.18
Ticino	0.01	0.07	0.25	1.20	2.53
Zurich	0	0.03	0.16	0.97	2.16

H^+-dominance vs 1st order dominance

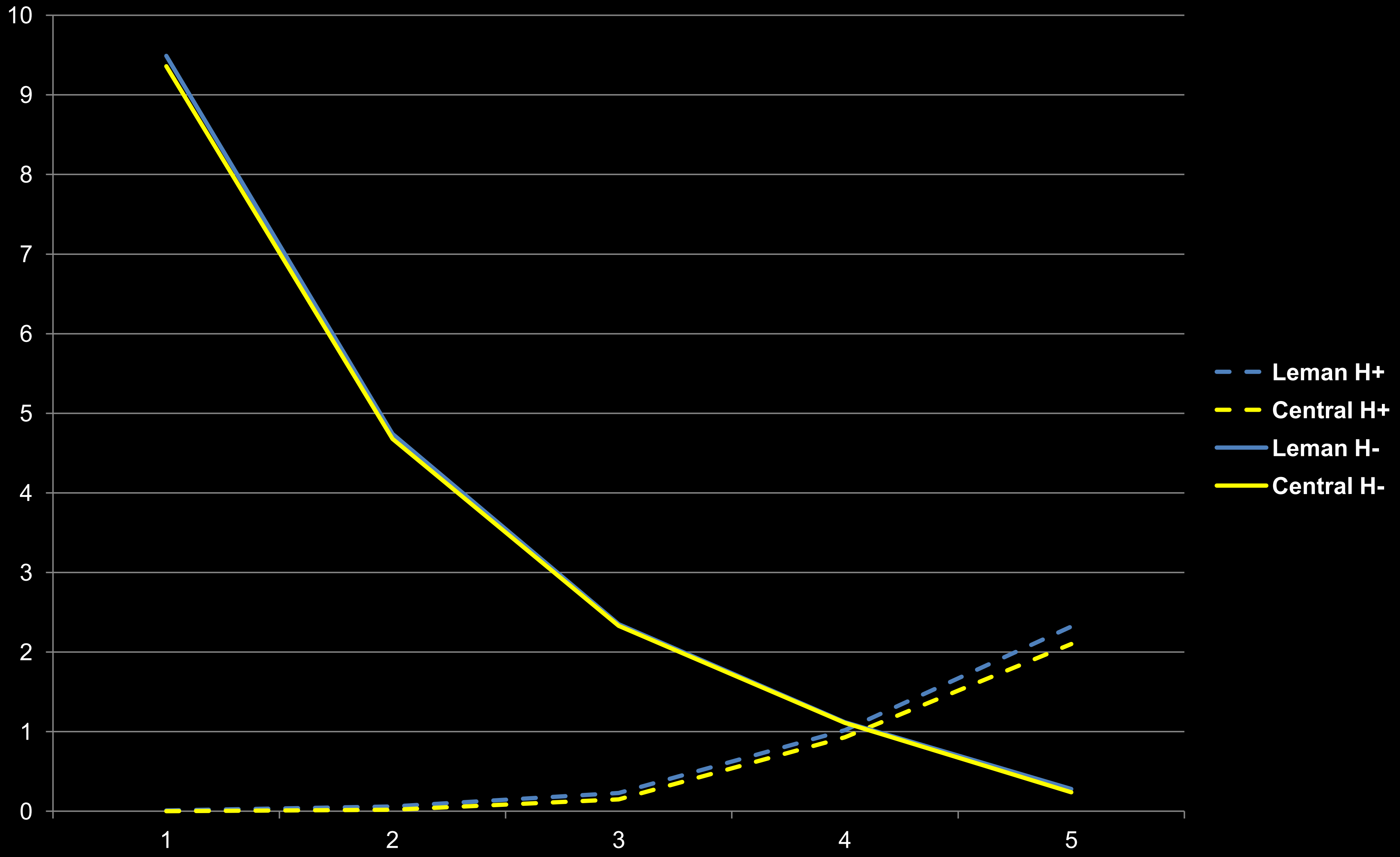
H^+-dominance



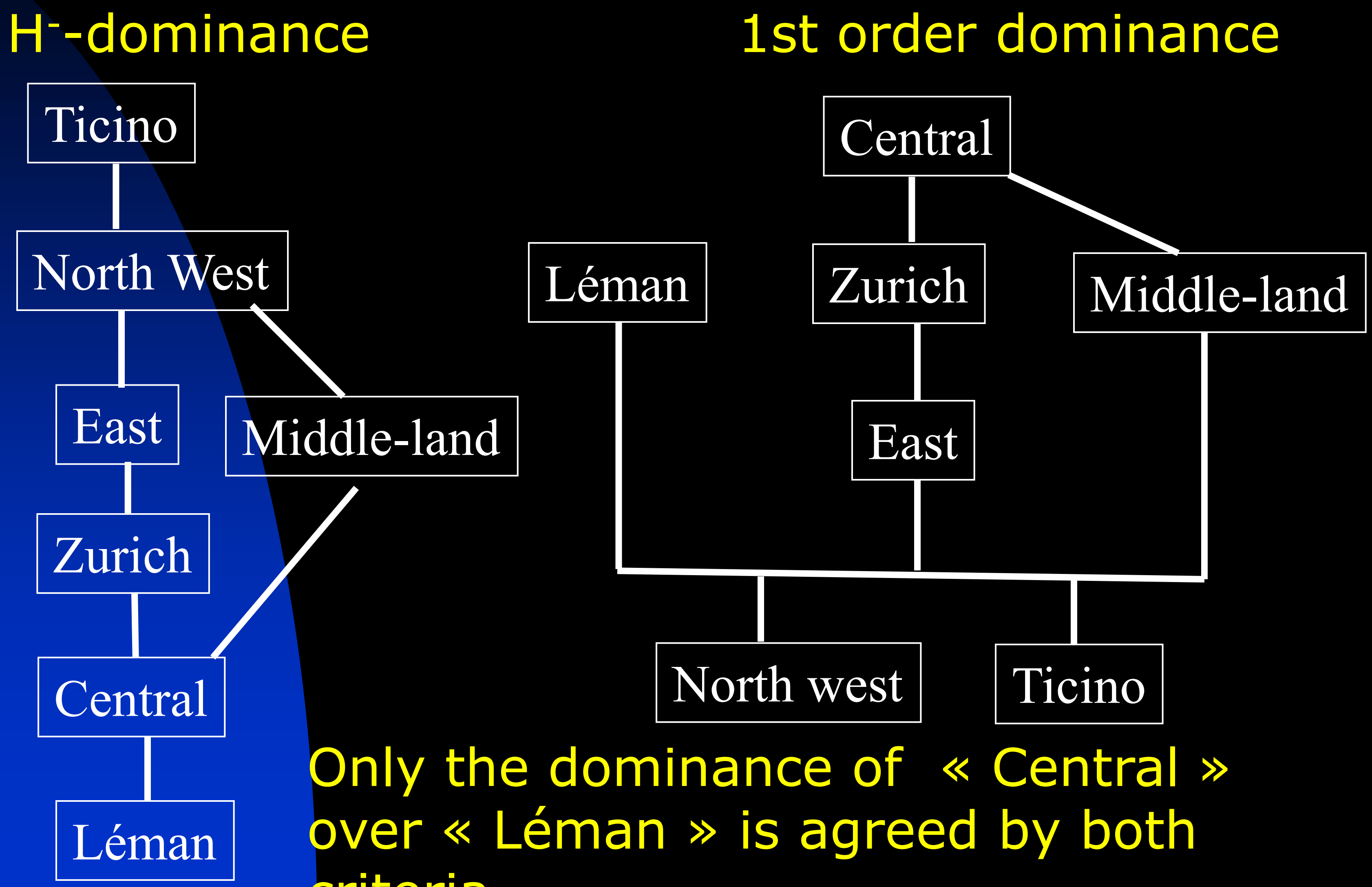
1st order dominance



Drawing these curves



H^--dominance vs 1st order dominance



Only the dominance of « Central » over « Léman » is agreed by both criteria

Comparing distributions of body mass index

- Obesity and overweight are increasingly recognized as major problems (both for health and for self-esteem)
- So can be « underweight »(anorexia).
- Body mass index (the weight in KG per squared meter of surface body) is often used as a diagnostic tool to identify threhold of pathologic weights:
- < 18 underweight (A)
- [18-25] norm (B)
- [25-30[over weight (C)
- [30-35[mild obesity (D)
- [35-40[severe obsesity (E)
- >40 Morbid obesity (F)

Comparing distributions of body mass index (2)

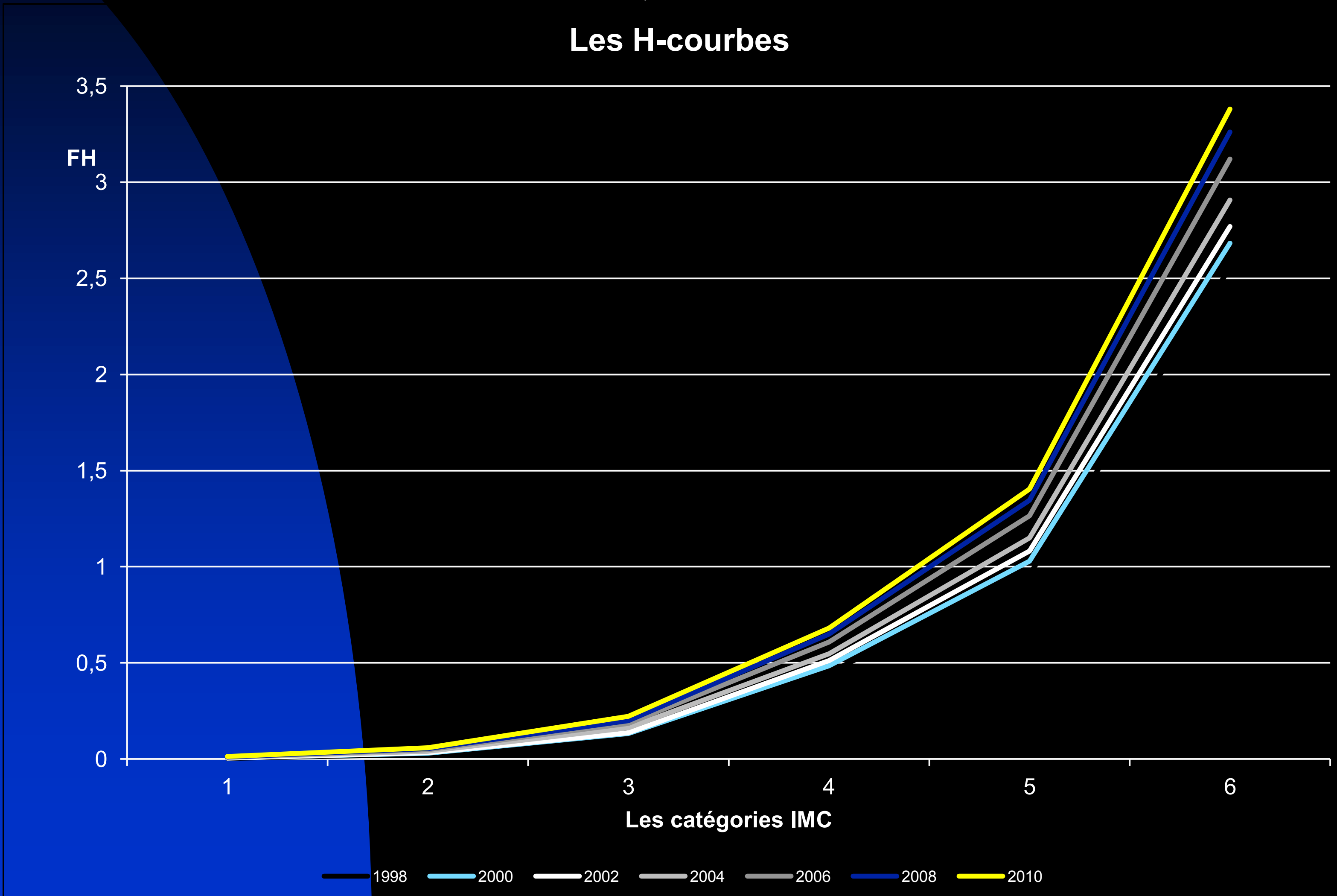
- Difficulty with BM: the ordering of the categories is not clear.
- It is clear that B $\succ$C $\succ$ D $\succ$E $\succ$ F. (1)
- But what about A ?
- It is clear that B $\succ$A (2)
- In the following we consider all possible rankings of those categories consistent with 1 and 2.
- We look at the evolution in France of the distribution of BM on a large sample of the French population (22 000 individuals, 8000 households).
- Data source: survey ESPS (panel since 1998)

Comparing distributions of body mass index (3)

- Except for the rankings of the categories for which D $\succ$ A, there is deterioration of the distribution of BM on the period 1998 – 2010 as per the H^{+} criterion for both French adult females and males.
- Except for the rankings of the categories for which A $\succ$ C, the distribution of BM among the French males H+ dominates the distribution of BM among French Females (in 2010).

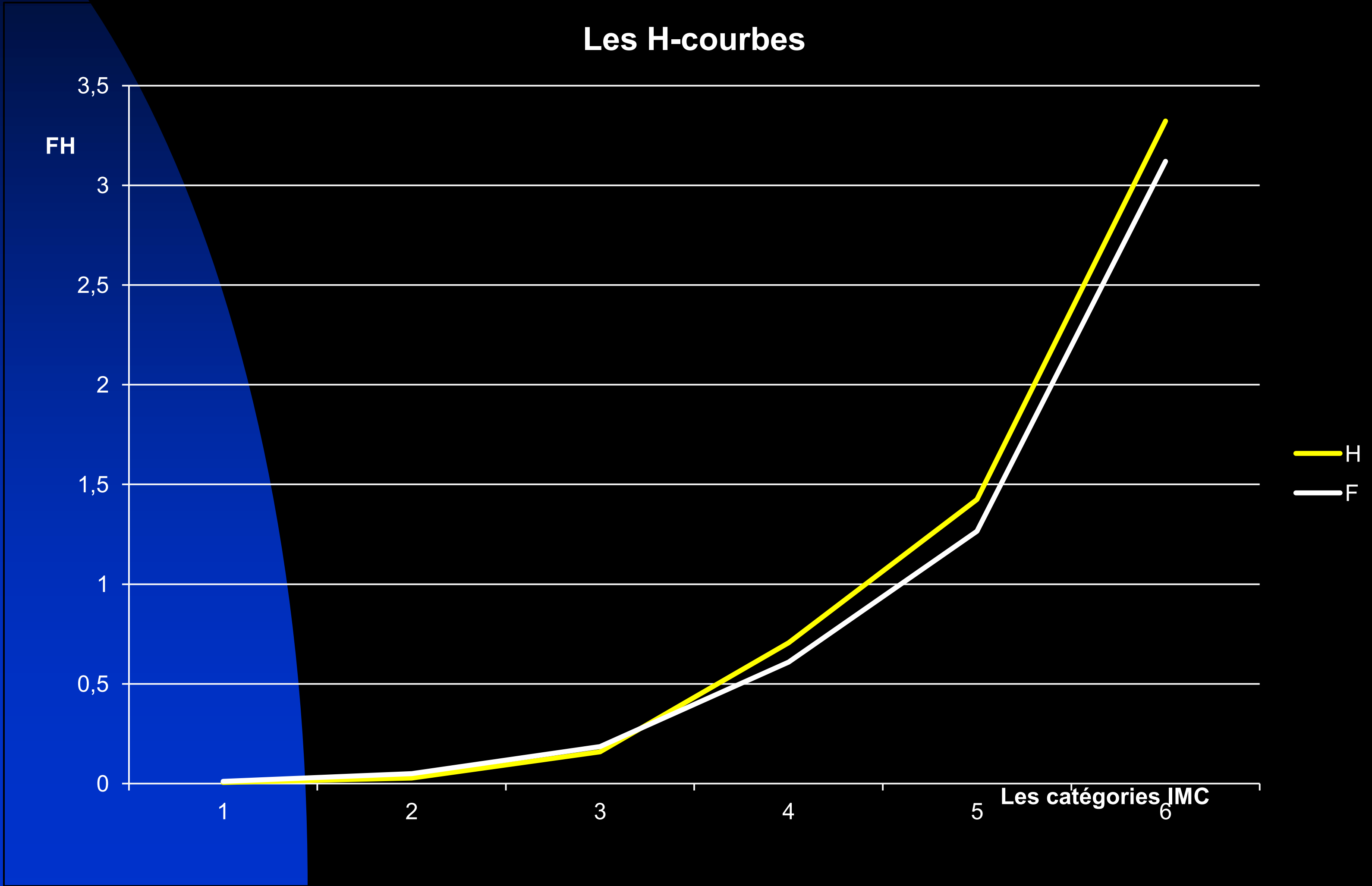
Some H+ curves

French adult females, $A \succ C$



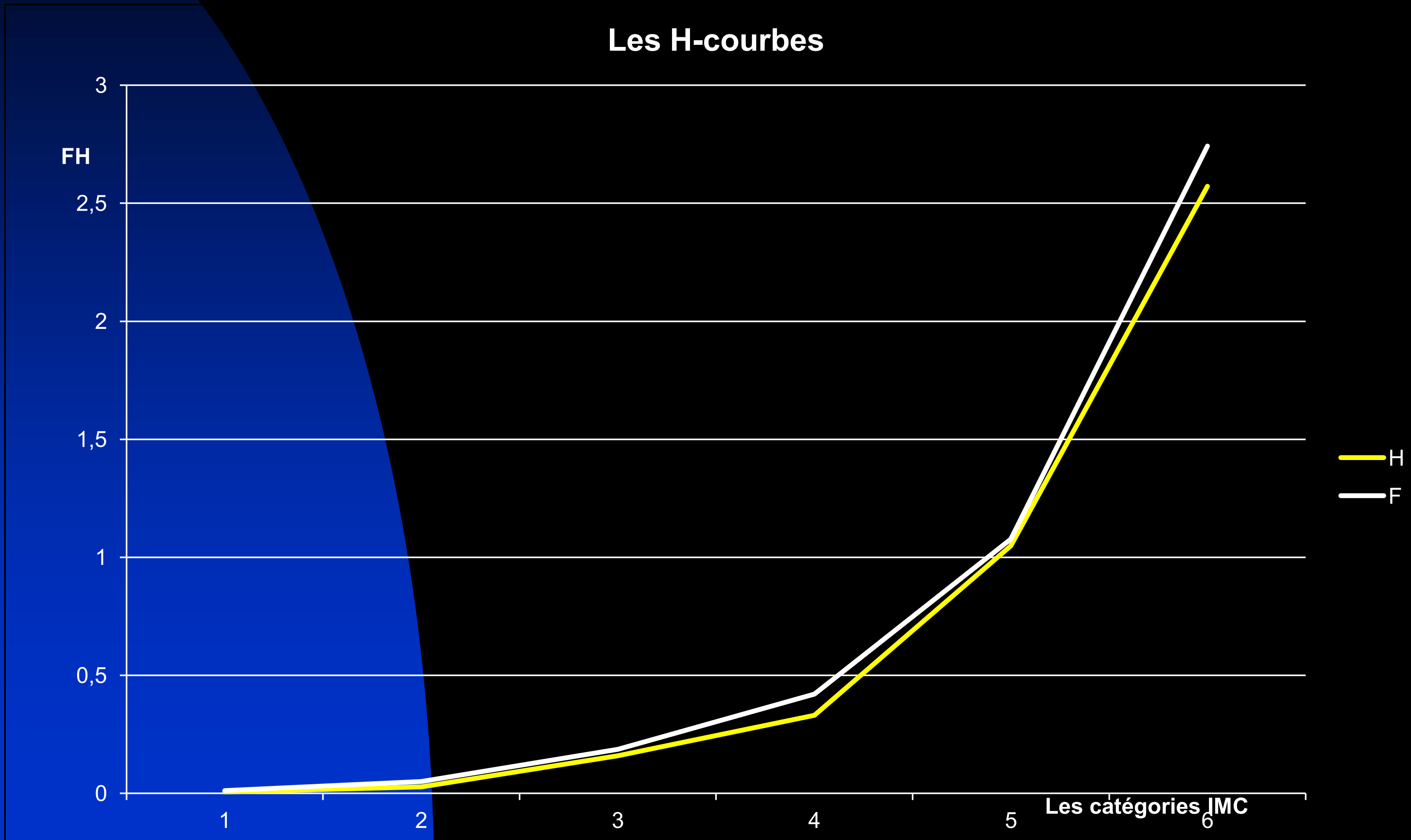
Some H+ curves

Male - Females, A $\succ$ C No dominance



Some H+ curves

Male - Females, C $\succ$ A $\succ$ D Dominance



Hammond transfers and classical social choice theory:

- In classical social choice theory, Hammond equity principle is somewhat tightly connected to the so-callled Leximin ordering.
- Ex: Bosman & Ooghe (2013 ?): the only continuous anonymous, Pareto-inclusive and Hammond-sensitive quasi –ordering is the maximin criterion.
- We don't fall into "leximin" trap due to the discrete nature of the scale.
- But suppose that we refine the grid of the ordinal attribute
- a t refinement: $\{1,\ldots,k\} \rightarrow \{1/2^t,\ldots.,2^t k/2^t\}$
- The H^+ criterion depends upon the grid (initial grid: $t=0$)
- $H^+(t)$ the definition of the H^+ -curve for refinement t (t=0,…)
- Theorem: there exists t at which $H^+(t)$ dominance and Leximin are equivalent.

Conclusion:

- We have provided a « foundational » theorem for normative evaluation dealing with distributions of a discrete ordinal attribute.
- Approach is easily workable
- Need to do:
- 1) develop ordinal inequality indices consistent with Hammond transfers,
- 2) Make empirical applications (Pisa Scores)
- 3) Multidimensional generalizations ???