Analyse & Kritik 23/2001 (© Lucius & Lucius, Stuttgart) p. 191–204

Marlies Ahlert/Hartmut Kliemt

A Lexicographic Decision Rule With Tolerances

The Example of Rule Choice in Organ Allocation

Abstract: The implementation of the Wujciak algorithm as a new rule for organ allocation by Eurotransplant is of considerable interest for the theorist of choice making. In the process reformers accepted the status quo in principle but expected that their potential opponents would be willing to make minimal or 'tolerable' concessions. Thereby the consensual introduction of new dimensions of value and reforms of allocation practices based thereupon became viable. The paper characterizes a decision procedure based on 'almost lexicographically pre-ordering established values and practices' in a stylized manner, presents a formal reconstruction of it and points out some of its potential implications for rule choices in general.

1. Introduction

Economists accept that in a world of scarce resources the cost of choosing an alternative is the best alternative forgone. Whenever allocating 'scarce means' that have 'alternative uses' to 'given ends' (see on this, of course, Robbins 1935) they suggest to focus on such 'opportunity costs'. Still when faced with so-called 'hard-choices' that concern life or limb of a specific human person, even some economists and certainly most other people resent to consider explicitly the opportunity costs of chosen alternatives. Making 'tragic choices' (see Calabresi/Bobbit 1978) can be avoided by relying on general rules which preprogram the decision of specific cases. The consideration of opportunity costs must not take place in the concrete application of the rule but only on the level of rule choice. And rules typically can be evaluated in terms of statistical effects on unnamed and personally unknown individuals. This is a considerable advantage if human lives are at stake (see Krämer 1982 for a general German account with respect to health care issues). In such cases we strongly prefer weighing statistical lives to weighing concrete lives.

Though the distinction between within rule choices (concerning concrete lives) and choices of rules (concerning statistical lives) is fairly standard (for a canonic account see Brennan/Buchanan 1985) it is not at all obvious which are the best methods of evaluating and making rule choices. Generating conceivable consensus on rule choices behind a veil of ignorance/uncertainty is often proposed here (on such veils see originally Vickrey 1948 and later on Buchanan/Tullock 1962; Harsanyi 1976; Rawls 1951). Though such a conceptual device is suitable for theoretical purposes it is not directly applicable to practical policy choices.

Therefore successful consensual policy change will in general have to rely on strategies other than appeals to some veil of ignorance/uncertainty notion.

Of course, consensus should be forthcoming whenever there are Pareto superior moves (see for canonic statements of unanimity in a Paretian framework Wicksell 1969; Buchanan 1999). But cases in which a policy advisor could reveal the existence of a Pareto improvement to a surprised audience seem extremely rare in the real world. If there are such moves on which all could immediately agree then only error and lack of information could have prevented those changes from emerging beforehand. So focus on Pareto improvements like reliance on some veil of ignorance/uncertainty will not help much in matters of practical politics. In this paper we shall try to keep reasonable close to both certain aspects of consensus buliding based on the Pareto principle and actual practices of consensual reform policies. As an alternative to consensus building based on the Pareto principle we suggest what we call a 'lexicographic decision rule with tolerances'. We will derive the general form of such a 'lexicographic decision rule with tolerances' from our stylized account of the particularly instructive specific real world example of the choice of the Wujciak algorithm by Eurotransplant. But before doing this some more general observations on the priviledged role of the status quo in section 2 seem appropriate.

Section 3 then sets the scene for the discussion of our paradigm example by presenting some features of the history of Eurotransplant that eventually led to the introduction of the Wujciak algorithm. Section 4 sketches for the purposes of the present discussion some of the data on which Thomas Wujciak based his central arguments along with some of these arguments. Then, in the central section 5, a model of choice making on the level of rule choices is presented and discussed in some of its formal features. Section 6 brings in the procedural aspect while section 7 contains some concluding observations.

2. Rational Conservativism?

The so-called 'rationalist constructivists' (see for a classical critique of their views Hayek 1973–79) of all times would just sneer at the argument that an established social practice simply because it is the established one has a strong argument in its favor. In particular, they would deny that in reforming and adapting our institutions the acceptance of path dependence could be rationally defended. But in the real as opposed to the theoretical world path dependence is important.

The contingent fact that the status quo does exist cannot be neglected. It holds as special place in our *actions* as well as in our *judgements*. First, any proposal in favor of a deliberate reform and improvement of what seems still to work reasonably well will be met with reluctance. The proverbial "never change a winning team" or "if it ain't broke, don't fix it" indicate this first common sense bias toward conservativism. Second, insisting on piece-meal reforms or that changes be small and reversible is not only a Popperian idea of quite some philosophical respectability it also corresponds to the common sense of the ages.

Most of the time we prefer to crawl rather than to jump in the presence of risk (see for fundamental accounts of some aspects of risk management Glickman/Gough 1990; Jasanoff 1986). Third, not only in action but also on the level of evaluation on which judgements are passed there exists a natural bias in favor of the status quo itself. Preference orders are not changed in view of very minor alterations along existing standards of evaluation. And even if completely new dimensions of value emerge common sense insists that value dimensions should not be changed arbitrarily.

We believe that the bias favoring the status quo is of general importance and that it is also very important to distinguish clearly here between the levels of evaluation and of taking action. Rather than reflecting on 'the status quo bias' in the abstract we intend to demonstrate its relevance and implications by a particularly instructive specific example. Reforming rules of organ allocation in Eurotransplant shows all the crucial features mentioned so far: potentially tragic or in any case hard choices, ascent to the level of rule choice, the necessity of consensual decision making about rule enactment, fundamental controversies over value issues, the problem of taking action or not and the potentially privileged role of the status quo.

3. Some Aspects of the Development of Eurotransplant

In the early years of organ transplantation organs were assigned in ways akin to assignment by lot. Who would receive an organ depended on who would have been where at what time. However, soon medical doctors learnt more about the mechanisms of HLA matching. It was then seen more clearly that it would be advantageous to have a large pools of potential recipients and of donated organs from which good matches could be picked.

For transplant-centers the need to gain access to a broader basis from which to choose in matching organs and recipients posed a fundamental challenge. The centers met this challenge by voluntary collaboration. Due to private initiative the institution of Eurotransplant emerged spontaneously as a non-government institution in which transplant-centers from Austria, the Benelux countries and Germany (since the year 2000 Slovenia is a member, too) co-operated.

Until the ascent of the modern immuno-suppressive drugs like cyclosporin finding good HLA matches was perceived as being of overwhelming importance. Improving the 'average' match between the characteristics of the tissue of transplanted organs and the tissue of the recipients as measured by the frequency distribution of the number of HLA mismatches was the aim of forming Eurotransplant. This aim was initially pursued by allocating one kidney centrally across the Eurotransplant region according to HLA match while leaving the other kidney basically with the center of ex-plantation.

If an organ was allocated centrally through Eurotransplant then the patient with the smallest number of HLA mismatches between his own and the graft's tissue would receive it. (The allocation was subject to certain other special priority rules like those favoring children and granting priority to realizing so-called full house transplants—transplants with zero mismatches, but these details may be put aside for the purposes at hand.)

Except for a potential increase of transaction-costs increasing the size of the organ pool is clearly conducive to the aim of improving average tissue match in organ allocation. But the policy of going for the best match at any point in time is too myopic to be optimal with respect to the overall allocation. It is quite obvious that the quality of HLA match in the emerging allocation could be improved by taking into account the comparative likelihood that a better organ might be found quite easily. More specifically: If for an organ that became available at a specific time the patient with the smallest number of mismatches on the waiting list could find an even better fitting organ with high probability soon while the chances of other patient(s) on the list to find an organ of comparable quality would be much lower then this should be factored in if the overall aim is to improve the quality of HLA matching across the whole set of transplantations.

Adopting less myopic allocation policies than the original one seems clearly desirable if improving 'average' HLA match is the only aim of central allocation policies. However, evaluating the quality of organ allocation exclusively according to HLA match did not remain unchallenged. With longer and longer waiting lists which emerged due to the scarcity of organs the number of individuals who were waiting for outrageously long time-spans while others received a transplant quite swiftly also increased. In fact, some individuals would wait merely six months for a kidney transplant while others were on the waiting list for more than eighteen years. Historically the criticism that outrageous discrepancies in waiting time existed was of the greatest importance. Statistically the differences between the expected survival time of a graft with, say one and two mismatches are not so strong that assigning the organ independently of all other value characteristics to the patient with the smaller number of mismatches would be self-evident. Though statistical data show that even today the expected survival time of the organ graft depends to some extent on the number of HLA mismatches between the tissue of donor and recipient (see Wujciak/Opelz 1993a; 1993b) it seems problematic to neglect differences in waiting time altogether.

In the early 90ties of the last century among those concerned with organ allocation at least two schools had developed. One insisted that organ allocation after making certain provisions for blood type, in particular should be evaluated exclusively according to the criterion of minimizing HLA mismatches. The standard of evaluation was basically the quality of tissue match that would emerge in the statistics of the results of organ allocation. This school could accept a forward looking allocation procedure increasing the quality of tissue match. But the school rejected a modification of the algorithm for reasons other than improving tissue match in the overall statistics.

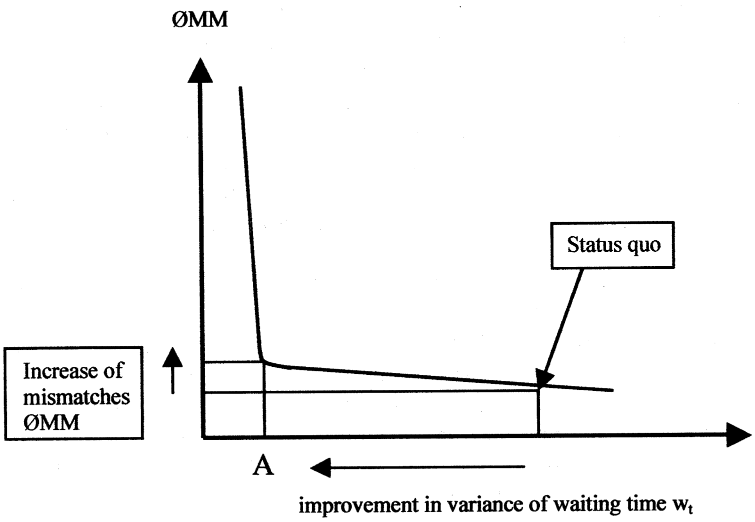
The other school pointed out that there were transplant-centers that allocated organs almost exclusively according to waiting time and would have basically the same medical success as measured in average, median or fifty percent survival times of grafts in the overall statistics of results according to which competing allocation procedures would be evaluated. So, the second school argued, there was no legitimate reason to leave out of account the distribution of waiting time altogether when evaluating rules for the allocation of organs.

Even doctors who favored HLA matching very strongly since they were convinced of its medical importance felt that extreme differences in waiting time were undesirable. Almost all were agreed that average and maximal waiting time of patients on the list as well as percentages of patients waiting more than three and more than five years, respectively, were important in evaluating the success of organ allocation. Under the established practices of Eurotransplant these characteristics had become increasingly unsatisfactory in the early 90ties already.

The tension between those who argued that a modification of the allocation procedure should bring about a more equitable distribution of waiting times in the transplant statistics and those who rejected any such measure grew. There were indeed serious concerns that this controversy in the end might threaten the sustainability of Eurotransplant as a voluntary organization altogether. It was feared that centers favoring waiting time might leave or found a new organization thereby reducing the size of the pool and thus decreasing the quality of the collective good created by intercenter collaboration.

4. A Stylized Account of Wujciak's Strategy

In the early 90ties Thomas Wujciak studied the data set characterizing the results of organ allocation under the then established Eurotransplant practices. 'Rational conservativism' (as explicated in section 2) clearly spoke against radical alterations of the practice. As evaluated in terms of expected survival time of the graft organ allocation according to HLA match led to acceptable results. True enough, the variance in waiting time was disturbing but it seemed very undesirable to threaten the whole system of allocation by sweeping changes. In view of this it was essential that Thomas Wujciak did not suggest that HLA matching should lose its central role. He rather showed that one could acknowledge the special status of the quality of tissue matching in *evaluating* allocation procedures and at the same time take into account the additional value dimension of waiting time within the limits of reasonable or tolerable sacrifices in the quality of HLA-matching. HLA would lose its position as the sole criterion of evaluation. But the historically established dominance of the aim of improving HLA matching would be respected by an almost lexicographic pre-ordering of this criterion in the envisioned multi-attribute evaluation of the allocation scheme. Extended simulations demonstrated that a significant decrease in the variance of waiting time and in the average number of patients waiting longer than three or five years, $\O w_t$, respectively could be realized at the expense of merely a very minor decrease in the average number of mismatches, $\O$MM. In such a low opportunity cost situation it was clearly reasonable to expect that adherents of HLA matching would tolerate the insignificant reduction in the quality of matches brought about by the suggested alteration.
The following figure 1 illustrates the low opportunity cost situation graphically.



Wujciak could opt for going all the way indicated by the horizontal arrow in improving quality of organ allocation as measured by reductions in the ('spread' of) waiting time measure $Øw_t$. At the same time this would not decrease the quality of organ allocation as measured by the average number of mismatches in the pool beyond a change deemed tolerable.

A large reduction in the fulfillment of the established criterion would have led to the immediate rejection of reform proposals. Wujciak was aware that established aspirations, fulfilled by the status quo, had to be respected by any proposal for improvement that he might make. His simulations of the consequences of alternative allocation rules as evaluated against the existing data set demonstrated that the advocates of change could be willing to respect the status quo in principle. Only very minor reductions along established value dimensions were necessary to allow for substantial gains along others.

It is indeed a widely accepted requirement of fairness that very minor reductions of one's own favored values even if they are the established ones be accepted if thereby the values favored by others can be furthered substantially. Since in a situation as illustrated in figure 1 the change away from the status quo is not required by a minor improvement but rather by a substantial one, fairness seems to dictate that the change be accepted.

In sum: *First, accepting the status quo as dominant implies that one does not go beyond minor tolerable reductions of status quo values of ØMM. Second, the alteration of the status quo is suggested for substantial gains as measured by $Øw_t$. Third, not neglecting unduly the values of others not yet established in the status quo requires to tolerate minor reductions provided that they allow for major increases in the fulfillment of the values pursued by others.*

The next table illustrates the concept of an almost lexicographic pre-ordering of one dimension of value with some tolerance for the inclusion of another dimension of value with several examples. All of the examples grant dominant status to the criterion of HLA matching. Except for the single attribute rule 'Eurotransplant' (former status quo rule) which would focus on HLA matching exclusively, all other rules include additional dimensions of value at 'tolerable opportunity costs' as measured in particular in terms of the emerging statistics of the quality of HLA matching. The entries in the table indicate how the alternative rules fare according to several value dimensions in Thomas Wujciak's simulations. It should be carefully distinguished here between the several components possibly showing up in the rules and the multi-attribute standards for evaluating the rules on which the present paper focuses. For instance, the quality of HLA matching as measured on the level of the overall statistics of allocation results can conceivably be brought about by different rules based on quite different allocation criteria. Conceivably the components of the algorithm actually used to allocate a specific organ to a specific recipient could be totally unrelated to the dimensions of value to be included in the evaluation of the statistical results. One could for instance endorse an algorithm creating a priority ranking among recipients according to, say, hair color, height and the size of the shoes. What mattered in Wujciak's simulations and the tables below is the overall statistical result of the allocation. However, it should be noted as well that as a matter of fact HLA matching was indeed considered as a component of the algorithm and so was waiting time etc. So there is a factual relationship between the standards of evaluation and the components of the rule of allocation. And certainly this relationship is helpful in rendering the allocation procedure acceptable to the general public. For the individual recipient will accept allocations much more readily if they are based on criteria that seem to have some direct intuitive appeal rather than relying exclusively on an indirect justification in terms of statistical consequences.

The rules in the next table start with the original allocation procedure 'Eurotransplant' basically relying on HLA matching. The other rules are derived from Eurotransplant as basic rule by including criteria other than HLA matching. XCOMB1 includes waiting time along with HLA matching. XCOMB2 further improves on waiting time but pays a price in terms of a more unbalanced exchange of organs across borders. XCOMB3 includes the national balance of organ exchange explicitly as an additional criterion of allocation yet at a minor cost in increased average waiting time and number of 1–2 mismatches.

Criteria	Eurotransplant	XCOMB 1	XCOMB 2 (No Balance)	XCOMB 3 (Regions Balance avg/Max)
0 MM	16%	25%	25%	25%
1–2 MM	48%	44%	46%	45%
3 MM	26%	24%	24%	24%
> 3 MM	10%	7%	5%	6%
Waiting time				
Avg / Max	2.1/20	1.7/12	1.6/12	1.7/12
>3/>5 yrs.	27%/11%	18%/4%	16%/3%	18%/4%
Balance				
Avg/max	15/50	8/28	17/95	4/19
G /B /NL/A		-42/43/-8/7	-120/79/32/19	-91/63/23/15
Distance				
Local	45%	46%	51%	46%
< 90 km	?	18%	19%	18%

Data as presented by T. Wujciak, 27.11.1995
Table 9 Cluster statistics on criteria of patient's value

Setting aside the further discussion of the specifics of the rules we can focus on the more general common pattern of an 'almost lexicographic ordering with tolerances'. In our next step we intend to sketch that general pattern in somewhat more formal terms.

5. The Model

Basic structure: $(A, C, X, \geq, d, \epsilon, \delta)$

where

A is the *choice set* containing the different alternative rules (e.g. Eurotransplant, XCOMB1, XCOMB2, XCOMB3).

C is a finite sequence of *criteria* (C_1, C_2, ..., C_n) deemed relevant in the case at hand (e.g. HLA-match, the conjunction of Ø and max waiting time, national balance of organ exchange for different countries). The sequence is ordered according to decreasing importance in the sense that earlier entries are at least as important as later ones. For the sake of simplicity we shall assume here that earlier entries are strictly more important than later ones. The order of importance may be determined by systematic or substantial argument. But it may also be the case and, as we indicated, quite characteristically so that the order of importance is set simply by historical precedent in the sense of "first in, first taken into account".

X is a sequence of variables (X_1, X_2, ..., X_n) measuring the extent to which the n criteria are fulfilled (e.g. distribution of HLA-MM).

$x = (x_1, x_2, \ldots, x_n), x' = (x'_1, x'_2, \ldots, x'_n)$ are vectors of realizations of variables (X_1, X_2, ..., X_n) as emergent under different choices of alternatives $a \in A$

evaluated according to the criteria C. For instance, in the case of the criterion "waiting time" we have to consider two components, average Ø and max waiting time. These pairs are rankordered according to the criterion by the decision-maker who also can assess distances between such pairs which we assume to be representable by some metric d.

$\geq = (\geq_1, \geq_2, \ldots, \geq_n)$ is a vector of transitive and complete quality rankings that order the realizations of each criterion; where these realizations themselves may be lists or vectors. These n orderings model the *perception* of the quality of a realization with respect to every single criterion.

d=(d_1, d_2 , ..., d_n) provides a vector of metrics defined on realizations of each criterion.

$\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ is a vector of 'tolerances'; where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n > 0$. Changes (especially losses) in the fulfillment of any criterion j that according to the metric d_j are smaller than ϵ_j are deemed insignificant. On the level of *evaluation* they cannot affect the ranking of an alternative in the preference order.

We assume that the perception of the quality of any realization with respect to each criterion is finer than the grid that the tolerances create. This means that for all $k \in 1, \ldots, n$ the following relation between ϵ_k and the ordering $\geq_k$ holds. If two realizations $x = (x_1, x_2, \ldots, x_n), x' = (x'_1, x'_2, \ldots, x'_n)$ are seen to be indifferent with respect to criterion k, i.e. $x \geq_k x' \wedge x' \geq_k x$, then $d_k(x_k, x'_k) \leq \epsilon_k$.

$\delta = (\delta_1, \delta_2, \ldots, \delta_n)$ is a vector of positive (non-negative) thresholds that must be transcended if a change in the realization value of any criterion is to be treated as a change sufficiently significant to warrant taking some *action*; where $\delta_1, \delta_2, \ldots, \delta_n \geq 0$.

Definition of a lexicographic ordering with tolerances

For each vector $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ of (evaluative) tolerances we have a lexicographic ordering with tolerances $\geq_{lex\epsilon}$

$$x \geq_{lex\epsilon} x' :\Leftrightarrow x >_{lex\epsilon} x' \vee x \sim_{lex\epsilon} x'$$

where

$$(x_1, x_2, \ldots, x_n) \sim_{lex\epsilon} (x'_1, x'_2, \ldots, x'_n) :\Leftrightarrow \forall i : d_i(x_i, x'_i) \leq \epsilon_i$$

The definition of $\sim_{lex\epsilon}$ introduces a relation between vectors by means of the metric defined on the components. For each dimension k there can be a different ϵ_k and a different appropriate metric d_k. Obviously, since the metrics d_k are symmetric the vector relationship so defined must be symmetric as well.

$$(x_1, x_2, \ldots, x_n) >_{lex\epsilon} (x'_1, x'_2, \ldots, x'_n)$$
$$:\Leftrightarrow \exists k \in \{1, \ldots, n\}[\forall i < k : d_i(x_i, x'_i) \leq \epsilon_i \wedge d_k(x_k, x'_k) > \epsilon_k \wedge (x_k >_k x'_k)]$$

The asymmetric part $>_{lex\epsilon}$ of $\geq_{lex\epsilon}$ is like the symmetric part a relation between vectors defined via the metric on the components. The first time the threshold

for a tolerance is transcended the evaluative ordering is fixed accordingly. Dimensions coming up later are irrelevant then. Obviously, the symmetric and the asymmetric part taken together disjunctively create a complete relationship.
We think that the concept of a lexicographic ordering with tolerances quite well captures the way in which we in practice pass value judgements. First, we evaluate the world along several dimensions. Second, as we argued in the opening sections of this paper there seems to be a natural proclivity to impose a kind of lexicographic structure on C, the list of value criteria that we apply. Moreover, aspiration levels formed within a contingent historical process will allow for consensual decision making on a collective level only if they are respected in a way akin to lexicographically preordering them.

Properties of lexicographic orderings with tolerances $\geq_{lex\epsilon}$

$\geq_{lex\epsilon}$ is - well-defined
- complete
- reflexive
- not transitive.

$\geq_{lex\epsilon}$ **is well-defined** since $\geq_k$ is an ordering for all k.

$\geq_{lex\epsilon}$ **is complete**, i.e. $\forall x, x' \in X : x \geq_{lex\epsilon} x' \vee x' \geq_{lex\epsilon} x$

To show completeness let $(x_1, x_2, \ldots, x_n)$ and $(x'_1, x'_2, \ldots, x'_n)$ be two realizations.
The following disjunction and its logical equivalent are trivial logical truths

$$[\forall k \in 1, \ldots, n : d_k(x_k, x'_k) \leq \epsilon_k] \vee \neg[\forall k \in \{1, \ldots, n\} : d_k(x_k, x'_k) \leq \epsilon_k] \Leftrightarrow$$
$$[\forall k \in 1, \ldots, n : d_k(x_k, x'_k) \leq \epsilon_k] \vee [\exists k \in \{1, \ldots, n\} : d_k(x_k, x'_k) \geq \epsilon_k]$$

Any x, x' will of necessity either fulfill the first (1) or the second (2) of the two alternatives in the preceding disjunction. In both cases this will imply $(x, x') \in \geq_{lex\epsilon}$ or $(x', x) \in \geq_{lex\epsilon}$ and thus completeness. Considering the two alternatives in turn demonstrates this:

(1) $[\forall k \in 1, \ldots, n : d_k(x_k, x'_k) \leq \epsilon_k] \Rightarrow x \sim_{lex\epsilon} x'$
$\Rightarrow (x, x'), (x', x) \in \geq_{lex\epsilon}$

(2) $[\exists k \in \{1, \ldots, n\} : d_k(x_k, x'_k) > \epsilon_k]$
Set $h := min\{k | d_k(x_k, x'_k) > \epsilon_k\}$
then $\forall k \in \{1, \ldots, n\} : k < h \Rightarrow d_k(x_k, x'_k) \leq \epsilon_k$
and $d_h(x_h, x'_h) > \epsilon_h$
Now, $d_h(x_h, x'_h) > \epsilon_h \Rightarrow x_h >_h x'_h \vee x_h <_h x'_h$
Since, except for the tolerance concept, the ordering is a lexicographic one the 'intolerable' difference along dimension h is sufficient to induce an

ordering of the full vectors. All dimensions, l, with $l < h$ show differences so small that values along these dimensions are treated as being equivalent. The dimension h either has $x_h >_h x'_h$ which is sufficient for $x \geq_{lex\epsilon} x'$ since h is lexicographically pre-ordered to all dimensions $m > h$ or dimension h has $x_h <_h x'_h$ which is for the same reasons sufficient for $x' \geq_{lex\epsilon} x$.

Thus we can conclude that $(x, x') \in \geq_{lex\epsilon}$ or $(x', x) \in \geq_{lex\epsilon}$.

$\geq_{lex\epsilon}$ is reflexive:

Consider $x = x'$. Then

$[\forall k \in \{1, \ldots, n\} : d_k(x_k, x'_k) = 0]$

$\Rightarrow x \sim_{lex\epsilon} x \Rightarrow x \geq_{lex\epsilon} x$

$\geq_{lex\epsilon}$ is in general not transitive:

Counterexample for n=2 criteria:

Let $\epsilon = (\epsilon_1, \epsilon_2) = (1, 1)$ be tolerances for changes (especially losses) in the criteria that are deemed insignificant.

Let $(x_1, x_2) = (1, 12)$
$(x'_1, x'_2) = (2, 8)$
$(x''_1, x''_2) = (3, 4)$

Obviously

$x >_{lex\epsilon} x'$ since the difference along dimension 1 is tolerable but the difference along dimension 2 is decisive in favor of x

$x' >_{lex\epsilon} x''$ since the difference along dimension 1 is tolerable but the difference along dimension 2 is decisive in favor of x'

$x'' >_{lex\epsilon} x$ since the difference along dimension 1 is intolerable and x'' fares better along the first dimension this is decisive in favor of x''

regardless of the second dimension.

As the example shows there is no guarantee that an almost lexicographic ordering with tolerances might not contain cycles. What happens in any sequential procedure of realizing improvements will be path dependent and show a high degree of sensitivity with respect to the precise timing and conditions of decision making. So let us turn to some issues of procedure.

6. The Procedural Aspect

Up to now we merely considered preference formation on the choice set A (i.e. in the case of Eurotransplant the set of alternative organ allocation rules from which to choose). Yet besides representation and evaluation of states of the world there is the possibility of intervening in the course of affairs such as to bring about more desirable states. If the situation that forms the status quo is characterized by x then there is a prima facie reason for action whenever there is a state of affairs x' with $x' >_{lex\epsilon} x$. But such an evaluation, though a prima facie reason for possible change, is not in general a sufficient reason for action. We do not instantaneously take action if there is a chance for improving our situation however slightly. The boundedly rational decision maker considers to give up the status quo in order to choose a new alternative only if she deems the improvement significant.

To capture this behavioral inertia of real world decision-making let us try to specify somewhat more precisely what the aforementioned statement that "a difference to make a difference must make a difference" might imply procedurally. If $x' >_{lex\epsilon} x$, significant differences in evaluative terms must prevail. At least the differences transcend the relevant thresholds. Whether these significant evaluative differences are behaviorally significant still depends. The list $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$ contains thresholds separating the behaviorally significant from the behaviorally insignificant. Only if these thresholds are transcended a significant change in the realization value of any criterion is to be treated as sufficiently significant to warrant action.

Relying on both, the evaluative and the action thresholds a plausible criterion for taking action to transform a state characterized by x into a state characterized by x' is:

$\exists k \in \{1, \ldots, n\}:$
$\forall i < k : (d_i(x_i, x'_i) \leq \epsilon_i$ – all differences prior to k are insignificant
$\lor x'_i \geq_i x_i)$ – or x' is not worse than x in prior criteria
$\land d_k(x_k, x'_k) > \epsilon_k$ – difference in k is *evaluatively* significant
$\land d_k(x_k, x'_k) > \delta_k$ – difference in k is big enough to trigger *action*
$x_k <_k x'_k$ – changing status quo is indeed an improvement

Consider again the previous example with $\epsilon = (\epsilon_1, \epsilon_2) = (1, 1)$ as tolerances for changes (especially losses) in the criteria that are deemed insignificant in terms of value. Let $\delta = (\delta_1, \delta_2) = (2, 2)$ be thresholds that must be transcended if a change in the realization value of any criterion is to be treated as sufficiently significant to warrant action.

Again (x_1, x_2) =(1, 12)
(x'_1, x'_2) =(2, 8)
(x''_1, x''_2) =(3, 4)

$x >_{lex\epsilon} x'$ and $x' >_{lex\epsilon} x''$ and $x'' >_{lex\epsilon} x$. The second dimension now shows significant differences in terms of evaluation as well as of action. For example if people would start with x'' they would switch to x' even though along the first dimension the difference is tolerable. Then comparing x' with x individuals would move to the state evaluated at x. The significantly different evaluation along the first dimension would lead to a cycle in evaluation but it does not any longer imply a cycle in action since $3 - 1 = 2$ is not greater than 2.

7. Concluding Remarks

The preceding argument is formulated in terms of a given list of criteria and fixed vectors ϵ, δ. What is deemed relevant as a dimension of value may itself be subject to change. The list C may, in principle, become longer or shorter, change by substituting an old by a new dimension or by permuting any two of the dimensions etc. It is also possible that what seems tolerable along prior dimensions is itself dependent on the criterion to be included in the list. Several more complicated interdependencies are conceivable and can be included without changing the basic structure of the lexicographic ordering with tolerances as presented in the model. Let it suffice here to give but one example by way of a final remark.

Let us assume that there is an established rule $a \in A$ that has been chosen according to the process described in our model and has been in place for a while. Then the argument from 'rational conservativism' applies to that rule too. Then the normative force of what is already established by the preceding history of an institutional rule restricts possible changes in the set of criteria to a lengthening of the list of criteria. Assume that the list for the existing rule is of length n. Assume also that it is suggested that criterion $n + 1$ be included in the decision making process. A change in the set of criteria that would lead to a different choice of a rule a′ from A is bound to induce some reductions in the degree to which the previous n criteria are fulfilled.

Again the issue of tolerable reductions naturally emerges. As opposed to the basic model let us assume that the vector of tolerances $1, \ldots, n$ is dependent on the perceived normative relevance of the criterion to be introduced and the possible improvements in the fulfillment of the criterion as brought about by possible actions. More specifically, $\epsilon_1(\delta_{n+1}), \ldots, \epsilon_n(\delta_{n+1})$. The latter list generates a class of possible consensual reforms of the established rule a. So if reformers can come up with a reform proposal such that fulfillment of criterion $n + 1$ will be improved by more than δ_{n+1} within the limits $\delta_1(\epsilon_{n+1}), \ldots, \delta_n(\epsilon_{n+1})$ then consensus should emerge that the criterion $n + 1$ should be included and that the new rule is better in view of the relevant trade offs.

Clearly the preceding sketch is quite close to the incrementalist trial and error process in which Thomas Wujciak himself has been looking for candidates a′ to 'beat' an established rule a of organ allocation. Incrementalism reminds us of the fact that the opportunity cost of a new rule is the old rule forgone.

Bibliography

Brennan, H. G./J. M. Buchanan (1985), *The Reason of Rules*. Cambridge
Buchanan, J. M. (1999), *The Logic of Constitutional Liberty*, Indianapolis
—/G. Tullock (1962), *The Calculus of Consent*, Ann Arbor
Calabresi, G./P. Bobbit (1978), *Tragic Choices. The Conflicts Society Confronts in the Allocation of Tragically Scarce Resources*, New York
Glickman, T. S./M. Gough (eds.) (1990), *Readings in Risk*, Washington D.C.
Harsanyi, J. C. (1976), *Essays on Ethics, Social Behaviour, and Scientific Explanation*, Dordrecht-Boston
Hayek, F. A. v. (1973–79), *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy*, 3Vols, London-Henley
Jasanoff, S. (1986), *Risk Management and Political Culture: A Comparative Study of Science in the Policy Context*, New York
Krämer, W. (1982), *Wer leben will muß zahlen. Die Kostenexplosion im Gesundheitswesen und ihre möglichen Auswirkungen*, Düsseldorf-Wien
Rawls, J. (1951), Outline of a Decision Procedure for Ethics, in: *Philosophical Review, 60*, 177–190
Robbins, L. (1935), *An Essay on the Nature and Significance of Economic Science*, London
Vickrey, W. (1948), Measuring Marginal Utility by Reactions to Risk, in: *Econometrica 13*, 319–333
Wicksell, K. (1969 (org. 1896)), *Finanztheoretische Untersuchungen. Nebst Darstellung und Kritik des Steuerwesens Schwedens*, Aalen
Wujciak, T./G. Opelz (1993a), A Proposal For Improved Cadaver Kidney Allocation, in: *Transplantation 56.6*, 1513–1517
— / — (1993b), Computer Analysis of Cadaver Kidney Allocation Procedures, in: *Transplantation 55.3*, 516–521