Econometrica, Vol. 67, No. 3 (May, 1999), 471–498

RELATIVE UTILITARIANISM

BY AMRITA DHILLON AND JEAN-FRANÇOIS MERTENS[1]

"If empirically meaningful interpersonal comparisons have to be based on indifference maps, as we have argued, then the Independence of Irrelevant Alternatives must be violated. The information which enables us to assert that individual A prefers x to y more strongly than B prefers y to x must be based on comparisons by A and B of x and y not only to each other but also to other alternatives."

Kenneth J. Arrow, "Social Choice and Individual Values," p. 112

In a framework of preferences over lotteries, we show that an axiom system consisting of weakened versions of Arrow's axioms has a unique solution. "Relative Utilitarianism" consists of first normalizing individual von Neumann-Morgenstern utilities between 0 and 1 and then summing them.

KEYWORDS: Axiomatization, social change, social welfare function, utilitarianism, welfarism, Arrow's Impossibility Theorem, expected utility.

1. INTRODUCTION

A SOCIAL WELFARE FUNCTION (SWF) maps profiles of individual preferences to a social preference. For preferences over lotteries, we axiomatize such a map, "relative utilitarianism" (RU), consisting of normalizing the nonconstant individual von Neumann-Morgenstern (VNM) utility functions to have infimum zero and supremum one, and taking the sum as social utility (Arrow (1963, Ch. III, §6, p. 32)).[2]

Our approach, in the sense of an axiomatic SWF, is very close to Arrow's tradition. The main difference seems to be the motivation. Given his insistence on the full strength of Independence of Irrelevant Alternatives (IIA), his approach seems more oriented towards understanding the voting paradox, and getting a general social choice paradox. Because voting situations are indeed characterized by successive votes between pairs, or at least small subsets of alternatives, imposing full strength IIA is almost necessary for analyzing the consistency between successive votes.[3] Our concern is more the normative

[1] This text presents research results of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. It was partly supported by NSF Grant #SES 8922610. The authors assume scientific responsibility. They are indebted to K. Arrow, C. d'Aspremont, and P. Hammond for comments.

[2] C.f. e.g., Rawls (1971, fn. 9, p. 22) for a very sketchy historical note on utilitarianism.

[3] Two other aspects also of Arrow's approach, which we share here, seem (in the current *zeitgeist*) more directed towards voting situations: one is the multi-profile formalism by itself, and the second consists in asking for a function rather than just a correspondence. In our view, however, they reflect exactly the original motivation of social choice theory, as a criterion for public policy recommendations. Indeed, to avoid complete "ad-hoc"-ness or even subjectivity, one has in such a position to be able to justify the way individual preferences will be taken into account; and "the" way means a function, while a justification must be under a "veil of ignorance," i.e., ignoring what those preferences happen to be.

question of finding a "good" SWF, effectively and justifiably usable for public policy recommendations. This is the original question of social choice theory, as well as the fundamental question of normative economics, and goes back implicitly at least two centuries ago, to the early utilitarians, and explicitly more than half a century ago, to the work of Bergson and Samuelson (Arrow (1963)). The full force of IIA is then no longer compelling: as discussed in point β, p. 481, some aspects of it are suspect, and contradict the most obvious intuition of what a good SWF should do. It is precisely those that we avoid.

While his axioms were inconsistent, ours yield a unique solution. IIA is substantially weakened, indeed almost dropped, its only remnants being the IRA axiom, the monotonicity axiom (MON),[4] and a continuity axiom. The other axioms are trivial cases of his Pareto axiom (or nontrivial cases of "Citizens' Sovereignty"), except for the classical anonymity axiom (ANON) strengthening his nondictatorship.

This paper clearly relies strongly on the additional mileage, in the form of more restricted preferences, obtained in decision theory by going to lotteries. There are many reasons for using this framework: First, "if conceptually we imagine a choice being made between two alternatives, we cannot exclude any probability distribution over those two choices as a possible alternative" (Arrow (1963, p. 20)). Another reason is methodological: in any completely symmetric situation, e.g. where the two individuals in a committee have opposite preferences on a set of two alternatives, the only choice of society preserving this symmetry is to randomize.[5] I.e., the most obvious axioms on social choice rules often force society to decide on a lottery. Further, *policy alternatives in social choice typically do involve very considerable risk and uncertainty*, at the level both of the individual and of society, such as the possibility of unemployment for the individual, or that of bad harvests and famines for whole nations. Acute debates on the differential *probability* of war under alternative policies (e.g., "unilateral disarmament") are not far away.[6]

Why not take interpersonally comparable utility functions as primitives.[7] The basic argument is Arrow's (1963), that primitives should be empirically meaning-

[4] Which can also be viewed as a weak analogue of his "Positive Association" (p. 96)—cf. Discussion p. 480.

[5] So total indifference *must* be allowed for society. For much deeper results on this theme, cf. Chichilnisky and Heal (1983). But if it were removed from the *domain*, a variant of our theorem would still hold (cf. previous versions of this paper, e.g. Dhillon and Mertens (1993, p. 47)).

[6] And clearly policy has to be sensitive to the individuals' attitudes towards such risks, notwithstanding Arrow (1963, p. 10). For example, with sufficient risk-aversion (for evidence, cf. Drèze (1981)), marginal returns from increasing a rich person's income may decrease so rapidly as to give a strong egalitarian flavor to the final outcome of RU.

[7] But cf. fn. 34 for why our results might no longer necessarily depend on this if we were to drop ANON.

ful[8] (e.g., p. 109, bottom; p. 112, top), lest the axioms themselves become meaningless, and so the whole theory. See also Rawls (1971, e.g., p. 322) for a more recent expression of the mythical character of any numbers behind preferences: they are just a construct in the "observing mathematician" 's mind, and without any uniqueness property in addition. But if until Arrow this ordinalist position was almost the consensus, apparently his theorem itself, together with the very influential work of Harsanyi, turned the tide partially, and led to the conclusion that interpersonal comparability was a must to obtain SWF's. The present theorem proves this conclusion false.

Even Harsanyi (176, p. 14) admits "the absence [for an individual] of an objective criterion ... for comparing his fellows' utilities with one another and with his own." Moreover, his justification of utility comparison by causal variables (1977, pp. 58–59), as well as his symmetry axiom, assumes utility, as opposed to just preferences, has an objective existence. In fact, he seems quite close to our own position: "The *metaphysical* problem (in interpersonal comparisons) would be present even if we tried to compare the utilities enjoyed by different persons with identical preferences and with identical expressive reactions to any situation. Even in this case, it would not be inconceivable that such persons ... should attach different utilities to identical situations, for, in principle, ... identical expressive reactions may well indicate different mental states with different people. At the same time, under these conditions this logical possibility of different susceptibilities to satisfaction would hardly be more than a metaphysical curiosity. If two objects or human beings show similar behavior in *all* their relevant aspects open to observation, the assumption of some unobservable hidden difference between them must be regarded as a completely gratuitous hypothesis and one contrary to sound scientific method. (This principle may be called the "principle of unwarranted differentiation." In the last analysis, it is on the basis of this principle that we ascribe mental states to other human beings at all: the denial of this principle would at once lead us to solipsism.) Thus in the case of persons with similar preferences and expressive reactions we are fully entitled to assume that they derive the same utilities from similar situations" (Harsanyi (1976, p. 15—footnotes omitted)).

Removing "and expressive reactions," this eloquent argument implies one should assign the same social preferences whenever the individual preference profile is the same, i.e. use a true SWF. And there are good reasons to remove it: expressive reactions that would not correspond to the expression of some preference are just as metaphysical. For those, Harsanyi's sentence can as well be turned around, to "different expressive reactions may well correspond to

[8] This is why, while formally our result applies to any interpretation of individual preferences, whether as tastes, as values, or as Harsanyi's "ethical preferences," our favored interpretation is that of "values," i.e., those that translate into actual choice behavior, e.g. when voting. Social preferences can only be determined by those that individuals *do* express.

identical mental states with different people": both sentences are just as true, or just as meaningless; what is indeed the meaning of identical or different mental states with different people (again interpersonal comparability—of mental states)? He concludes (1977, p. 58) as we do: "If all individuals' personal preferences were identical, then we could ascribe the same utility function U to all individuals." So presumably if two individuals' preferences coincide, we can assume their utilities do....[9]

The use of the sure-thing principle for society has been criticized by Diamond (1967); cf. Harsanyi (1976, Ch. V, also Ch. IV) and Sen (1970, p. 145) for a discussion. Further, if Harsanyi's arguments were not convincing enough, the criticism apparently relies on the idea that the individuals of a two-person society themselves strictly prefer a procedure of society consisting of choosing at random between $(1,0)$ and $(0,1)$ to a fair gamble (of nature) between two societies, one which would give them everything and the other nothing. In a sense, they have a positive disutility per se for being "unjustly" treated by society, in addition to their utilities for the final outcomes: their preferences depend on the procedures, and not just on the consequences. Our theorem remains applicable, taking the procedures as the set of alternatives.[10]

A widespread misunderstanding of Harsanyi's theorem (1976, Ch. II) consists of believing that, by multiplying each individual's utility function by the corresponding scaling factor, or by using some form of symmetry, one obtains the sum of individual utilities as social utility. With this in mind, one is then led to the impression that the present paper adds very little: just a rather heavy axiomatization for a specific normalization. While formally correct, this interpretation hides completely the fact that an individual's scaling factor can depend on the whole preference profile, cf. the example on MON p. 485 (which satisfies complete anonymity and neutrality). The "individual utility functions" resulting

[9] A slightly more delicate argument would be: if "intensities" were to mean something, the only possible precise interpretation of their *intrapersonal* comparison is that p_H is more preferred to q_H than p_T is to q_T if, Heads and Tails being equally likely, the combination (p_H, q_T) is preferred to (q_H, p_T) (Harsanyi (1977, Ch. IV, fn. 3; or 1976, p. 49)). So if someone's preferences are unchanged, his intensities are too, except possibly for being all multiplied by the same factor; he can only claim to have become himself more sensitive, irrespective of the specific alternatives under consideration, and only relative to the others: if everyone doubles his sensitivity, nothing changes. This is the claim of Sen's brahmin (1973, pp. 81–85): to deserve a bigger share of the cake, being a more efficient pleasure machine. It is one no individual can meaningfully make; such questionable judgements, as to the relative intensity of different individuals' pleasures, belong to Harsanyi's "moral observer," and have no place in our framework based purely on individual values. It is Harsanyi's "metaphysical curiosity" in its purest form.

[10] But if aggregating inconsistent preferences thus led society to prefer strictly a random choice between $(0,1)$ and $(1,0)$ to each of those pure alternatives, there may be another pure alternative, say $(0.6, 0.6)$, which society strictly prefers to each of the two, but prefers strictly less than the random choice. According to such preferences, society would select the latter, leading with *certainty* to an outcome where one will regret the foregone alternative $(0.6, 0.6)$: if individuals' preferences were really so, a strong case could be made that, rather than exposing society to such inconsistencies, individual preferences should first be corrected, and made to depend only on the consequences (Hammond (1988)), before being used in the SWF.

from such a multiplication become then arbitrarily complex functions of the preferences of all *other* individuals, making the sum-formula basically meaningless. The main purpose of the present axiomatization is to obtain the required separability: each individual's utility function in the sum-representation is independent of the others' preferences. While Harsanyi's Impartial Observer Theorem (1977) is based on the notion of a "moral observer" who, before being born, considers himself equally likely to be born as person 1 or person 2 and, assessing (with state-dependent utilities) his preferences between the two pairs of prospects (A, B) and (A', B') for persons $(1, 2)$ obtains a full profile of VNM utilities with comparable scales such that his ex-ante utility for an alternative ("act") is the average of the individual utilities for it.[11] Again, the moral observer fixes in the model the scaling of each individual's utility function; hence this scaling is left as a degree of freedom for the *user* of the theory to reflect his own ethical views (Harsanyi (1977, Ch. 4)). So this approach brings us back to square one: How does one aggregate the different individuals' "ethical preferences," i.e., those as a moral observer? This is then the question addressed here.

"Welfarism," on the other hand (e.g., d'Aspremont (1985), or Moulin (1988 Part I)), has interpersonally comparable utility functions as primitives, i.e., the *user's* assignment of individual utility functions to individual preferences, and so his own views on the "right" interpersonal comparisons.[12]

Observe that our approach leaves at first sight no scope for further ethical judgement in the use of the SWF. Indeed, the primitives are just individual preferences rather than explicit utility functions that would be associated with them so as to reflect the user's moral judgement. And the axioms determine the

[11] Harsanyi (1992, p. 681) even claims the moral observer can compare 1 in state A with 2 in state B, but this is superfluous.

[12] So in each case the user is in principle faced with the impossible task of empirical psychology to ascertain *each* individual's utility-intensity, solely by observing his expressive reactions. This impossibility means that in fact those weights are left at the user's discretion; he should thus assign them using his own value-judgements. So: (a) one contradicts one of the very foundations of the theory used; (b) since the user is not going effectively through his psychological exercise for each individual, he uses in effect only a map from preference profiles to a social preference, i.e. a SWF; (c) the only theoretical restriction kept on this function is the Pareto axiom, leaving thus to the user, through his choice of weights, any arbitrary selection from the Pareto correspondence to express his "values." This is dictatorship by the user: no need for any theory, he could (and should as said above) just as well dictate the end-result he wants (e.g., by assigning sufficiently high weights to preferences like his own), irrespective of the preferences of any other individual. Finally (d), this total discretion is left to the user through the use of a completely unwieldy number of parameters, whose relation to any explicit social philosophy is very indirect, and completely obscured by the interplay with variations in the preference profile. The end-result is that the whole scheme is self-defeating, and so obviously both surrealistic and ill-founded that the only effective uses of this theory we know of refrain completely from using such weights to express some form of ethical judgement, and start quite understandably by assuming, not only that all individuals with the same preferences have the same utilities, i.e. a SWF, but even (to avoid having to assign different weights to different preferences) that all have the same preferences, "hence" the same utilities.

Our framework allows one to take fully into account the true distribution of individual preferences, as well as an explicit philosophy of the state (cf. infra), parameterized (by the universal set $\mathcal{A}$) completely independently of individual preferences.

map uniquely, again removing any scope for ethical judgement in the choice of it. It could be argued that the ethical judgement is therefore completely explicit in our axioms. However, no room being left, in the choice of the SWF, for any judgement based on the circumstances of the case, we would not be comfortable with this situation, and in this respect might even to some extent have preferred to have axiomatized a small and relatively flexible family of SWF's—*if judgement, and thus flexibility, did not enter the final result another way*. Indeed, the final result is strongly influenced by the choice of the set of alternatives A, in fact, just by the best and the worst alternative for every individual. Besides the obvious feasibility restriction on A, the axioms themselves clearly imply that A is also limited by justice (including the constitution, the philosophy of the state, the social contract).[13] Individualism e.g. clearly implies there are no wrong alternatives in A, since they are all indifferent when individuals are indifferent among them.

This formulates the question much more operationally. Instead of needing as input the judgement of a mythical moral observer as to the different individuals' intensities of feelings, or something akin, with no empirical counterpart nor unambiguous or objective meaning, the ethical input here is the set of alternatives available to society. The ethical debate acquires thus a much more operational meaning, and one much closer to the traditional questions of social philosophy and ethics, which is where it belongs.

In particular, to apply RU meaningfully, one has to, and it suffices to, consider a set of alternatives sufficiently encompassing as to include, besides the actual alternatives of interest, each person's best and worst alternative within the "universal" set A, limited only by feasibility and justice. This leads in turn to a concept of "absolute utility:" the "correct" scaling of an individual's utility is determined solely by his own preferences and by the philosophy of the state adopted. This brings us almost back to classical utilitarianism, but this time, without any "moral judgement" as primitive, and with a complete axiomatization.[14]

For two alternatives, RU amounts to majority rule, so our result can be viewed as a generalization of May (1952)[15] that, for this case, majority rule is the only "reasonable" solution. And, when viewed as a mechanism, RU suggests letting each voter assign to every alternative some utility in $[0, 1]$, and to choose the alternative with the highest sum. Except possibly with very small sets of voters, voters will clearly find that, for their vote to have a maximal effect, they should assign either 0 or 1 to every alternative. Hence the corresponding direct mechanism seems to be "approval voting" (Brams and Fishburn (1978)).

[13] Traditional ethics' categories are good and bad, not to maximize some "goodness" function as in utilitarianism.

[14] Though our title is a clear reference to Kalaï and Smorodinsky (1975), possibly Bentham himself might recognize RU as the proper interpretation of "Everybody to count for one, nobody for more than one," his version of "one man, one vote."

[15] It shares with May's theorem its refusal of interpersonal comparisons as primitives. Would a claim that society ought to adopt the minority's view, just on the grounds this claims "stronger feelings about it," not be rejected?

To summarize, we consider only preferences as empirically meaningful primitives, and hence axiomatize *their* aggregation—coming thus back to the original formulation of social choice theory (Arrow (1963)), as a guide to policy recommendations, not as the ultimate foundations of ethics. So a priori the question of interpersonal comparisons is not even meaningful in our framework. Following this approach requires that the set A consist of all feasible and just alternatives: the ethical question lies in the choice of A and it is this that will determine whatever (implicit) "interpersonal comparisons" occur. The questions of justice being accounted for by the choice of A, the choice within A is to be handled by an appropriate generalization of majority voting: the SWF.

Section 2 introduces the axioms. Section 4 contains the main theorem, and Section 6 deals with the two exceptional cases not covered there. Section 5 shows each axiom is needed, even with the others interpreted in their strongest form. The idea of the theorem is simple, and its heuristics are presented in Section 3: in a Pareto representation, the individual weights are in general functions of the whole profile. The monotonicity axiom implies then the required separability, that each individual's weight depends only on his own preferences. IRA implies their consistency, and finally anonymity yields their equality across individuals. Such a proof can in fact be pursued with slightly stronger forms of MON, but this would not preserve the necessity results in Section 5. So in fact we have to prove all properties together, and need all axioms at every step of the proof.

2. THE AXIOMS

The following axioms, as well as the definition of RU on p. 471, will be assumed throughout from Sect. 3 onwards.

FRAMEWORK: The sets A of alternatives and N of individuals are fixed, and $\#N < \infty$. The set of lotteries Δ^A consists of all purely atomic probabilities on A, or of those with finite support. Preferences on Δ^A follow any standard axiomatization, ensuring a representation by affine functionals. This typically involves a continuity axiom; we explicitly assume, in addition if necessary, *individual* preferences to satisfy the following strengthened continuity axiom: $p \nsim q \Rightarrow \exists \varepsilon > 0$: $\forall r, (1-\varepsilon)p + \varepsilon r \nsim q$. Equivalently, the affine functionals are bounded, hence (duality (l_1, l_∞)) represented by bounded real-valued utility functions on A.[16]

INDIVIDUALISM AXIOM (INDIV): *If all individuals are totally indifferent, so is society.*

[16] Recall our set A consists only of those alternatives that are both feasible and just. So even if for mathematical convenience one were to consider utility functions say of income that are unbounded over $\mathbb{R}_+$, on the subset A they would still clearly be bounded: feasibility bounds them upwards, and justice downwards.

NONTRIVIALITY AXIOM (NONT): *The SWF is not constantly totally indifferent.*

NO-ILLWILL AXIOM (NOILL): *It is not true that whenever all but one individual are totally indifferent society's preferences are opposite to his.*

ANONYMITY AXIOM (ANON): *A permutation of individuals leaves social preferences unchanged.*

As to continuity,[17] we first discuss it in case A is finite.

Let $\mathscr{T}$ be the topology on preferences with the sets $\{\preccurlyeq \mid p \prec q\}$ as open sub-basis. For A finite, it is the strongest for which the mapping from utility functions (viewed as points in $\mathbb{R}^A$) to preferences is continuous (i.e., a quotient topology). By normalizing utility functions, the space can be viewed as a sphere plus one point (total indifference), with as open sets those of the sphere, and the whole space.

Then we would require the following (separate) continuity property: Assume $(\preccurlyeq_n^r)_{n \in N}$ is a sequence of preference profiles such that for some $n_0 \in N$, $\preccurlyeq_{n_0}^r$ converges to $\preccurlyeq_{n_0}^\infty$, which is not complete indifference, and $\preccurlyeq_n^r = \preccurlyeq_n^\infty \forall n \neq n_0$. Then $\preccurlyeq^\infty$ is a limit point of $\preccurlyeq^r$ (i.e., either the social preference at the limit is total indifference, or its normalized utility representation is the limit of those along the sequence).

We want to weaken this continuity property, and to define it very directly in terms of preferences, without using rather arbitrary topologies on preferences for infinite A. So we first weaken it to a closed graph property,[18] then take the strongest possible convergence for preferences: convergence along "straight lines" (even $u_n(a)$ is fixed except for $n = n_0, a = a_0$), where all (Hausdorff vector) topologies coincide:

DEFINITION 1: A *sequence of preferences* $\preccurlyeq^r$ *converges specially to* $\preccurlyeq^\infty$ if $\exists a_0 \in A$ such that $\forall r < \infty$, $\preccurlyeq^r = \preccurlyeq^\infty$ on $\Delta^{C\{a_0\}}$ and $\exists p^r \in \Delta^{C\{a_0\}} : a_0 \sim^r p^r$, and if $p \succ^\infty q \Rightarrow \exists r_0 : p \succ^r q \, \forall r > r_0$.[19,20]

[17] At least when the set A of alternatives is finite and under slightly stronger forms of MON (e.g. (cf. infra) "goodwill," or "consistency" and $S_u \neq 0$), RU follows generically (e.g., whenever any three individuals' utilities are independent) already without continuity. (Indeed, the proof of Lemma 1 goes through without it when knowing from the outset that the λ's are nonzero, and for part V that $\#A < \infty$, and the induction that follows is trivial when considering only such profiles.)

Still, we felt it should be kept, because (a) SWF's like those in example CONT, p. 485, are too close to election mechanisms to be dismissed by a mere nongenericity argument, and (b) the axiom is really needed when $\#A = \infty$, to ensure that utility functions that achieve their bounds have the same weights as the others.

[18] I.e., requiring $f(x_n)$ to converge to $f(x)$ only when, in addition to $x_n \to x$, one also knows that $f(x_n)$ converges.

[19] Replacing the last clause by "where the sequence p^r is monotone in the fixed preferences on $\Delta^{C\{a_0\}}$—and if, in case p^r is nondecreasing, $q \succcurlyeq^\infty p^r \forall r \Leftrightarrow q \succcurlyeq^\infty a_0$, and dually if p^r is nonincreasing" leads to an equivalent axiom, phrased even more directly in terms of preferences.

[20] Any constant sequence converges specially, and any specially convergent sequence has a unique special limit.

CONTINUITY AXIOM (CONT): *Let* $(\preccurlyeq_n^r)_{n \in N}$ *be a sequence of preference profiles such that for some* $n_0 \in N$ $\preccurlyeq_{n_0}^r$ *converges specially to* $\preccurlyeq_{n_0}^\infty$ *and for* $n \neq n_0$ $\preccurlyeq_n^r = \preccurlyeq_n^\infty$. *If the corresponding social preferences* $\preccurlyeq^r$ *converge specially, then* $\preccurlyeq^\infty$ *is a* $\mathcal{F}$*-limit point.*[21,22]

INDEPENDENCE OF REDUNDANT ALTERNATIVES AXIOM (IRA): *If 2 profiles, such that any lottery is unanimously indifferent to one in* $\Delta^{A'}$ $(A' \subset A)$, *coincide on* $\Delta^{A'}$, *their social preferences do so too.*

IRA restricts IIA to the case where both before and after the change the "irrelevant" alternatives are in essence lotteries on other alternatives: deleting them leaves the whole bargaining situation unchanged. It is so innocuous that a *much* stronger[23] form, taking the set of feasibility utility-vectors as primitive, is taken for granted in the formulation itself of the bargaining problem (which differs only in the additional datum of a "disagreement" point) without ever having been challenged.

MONOTONICITY AXIOM (MON): *Let* $\preccurlyeq$ *and* $\preccurlyeq^*$ *denote the social preferences for the profiles* $(\preccurlyeq_n)_{n \in N}$ *and* $(\preccurlyeq_n^*)_{n \in N}$, *where* $\preccurlyeq_i^* = \preccurlyeq_i$ $\forall i \neq n$ *and* $\preccurlyeq_n^*$ *is complete indifference. For any 3 lotteries* p, q, r: $r \sim^* p \succcurlyeq^* q \,\&\, p \sim_n q \succ_n r \Rightarrow [q \sim r \Rightarrow p \succcurlyeq q]$*—and the same holds with strict instead of weak preferences.*[24]

To understand MON,[25] consider e.g. the following "goodwill" axiom:[26]

$$p \sim^* q \,\&\, p \succ_n q \Rightarrow p \succ q.$$

Applying it to the pair (p, r) immediately implies MON. The following very weak Pareto-like (in terms of n's preferences and those of the rest of society) axiom does so too, comparing p with $\frac{1}{2}(q + r)$ (since $q \sim r$):

$$p \succcurlyeq^* q \,\&\, p \succ_n q \Rightarrow p \succcurlyeq q \text{—and the same with strict inequalities.}$$

[21] I.e., $\preccurlyeq^\infty$ is either total indifference, or the special limit of $\preccurlyeq^r$.

[22] To relate completely our axiom system to Arrow's, note that CONT too is an implication of IIA: this yields that, if $p \not\sim_{n_0}^\infty q$, then $p \succ^\infty q \Rightarrow \exists r_0$: $p \succ^r q \,\forall r > r_0$. CONT being trivially true if $\preccurlyeq_{n_0}^\infty$ is total indifference, the above implies the desired conclusion whenever the limit of $\simeq Ur$ is not total indifference, and otherwise $\preccurlyeq^r$ equals total indifference $\forall r$, so the conclusion follows too. In a sense, CONT relaxes IIA to hold only approximately, and only locally (i.e., when sufficiently many "irrelevant" pairs also compare in the same way—cf. quotation p. 471, and point β p. 481).

[23] Implying e.g. neutrality (and being its "correct" formulation for lotteries: "the different prospects only matter by the preference relations among them"). Even our full axiom system allows non-neutrality on the "exceptional profiles."

[24] One could *equivalently* replace the clause $p \sim_n q$ by $p \succ_n q$, for greater (?) plausibility.

[25] The Axiom is *so* weak that any weaker unquantified statement involving only two lotteries p and q would be trivial. A fortiori it cannot be axiomatized in such terms, and the above is the simplest possible axiomatization involving no quantifiers over lotteries. This is why we now relate it to more familiar conditions involving only pairs of lotteries.

[26] Just $p \sim^* q \,\&\, p \succ_n q \Rightarrow p \succcurlyeq q$, together with $p \sim q \,\&\, p \sim_n q \Rightarrow p \sim^* q$, also implies MON.

Also, the following "consistency" axiom[27] does:

$$p \succcurlyeq^* q \,\&\, p \sim_n q \Rightarrow p \succcurlyeq q \text{—and the same with strict inequalities.}$$

So any of the above mentioned requirements would already imply the axiom,[28] and the requirement used can vary both with the profile and with the triplet (p, q, r).

Notation

For any profile $\vec{u}$ the feasible set $R_{\vec{u}} = \{(u_n(p))_{n \in N} \mid p \in \Delta^A\}$ and $d_{\vec{u}} = \dim R_{\vec{u}}$. Henceforth utility functions are viewed as being in the quotient of the affine functions on Δ^A by the constant functions.[29]

PROPOSITION 1: *Let S, S^* and u_n represent $\preccurlyeq$, $\preccurlyeq^*$ and $\preccurlyeq_n$. Conditions* 1–4 *are each equivalent to MON:*[30]

1. $\exists \alpha, \mu, \tau$: $\alpha S + \mu u_n + \tau S^* = 0$, $(\alpha\mu, \mu\tau, \tau\alpha) \neq (0,0,0)$, $\min(\alpha\mu, \mu\tau, \tau\alpha) \leq 0$;
2. *either* $S = \pm S^*$ *or* $u_n = \alpha S + \tau S^*$, $\max(\alpha, \tau) \geq 0$;
3. *either* $S^* = \pm u_n$ *or* $S = \mu u_n + \tau S^*$, $\max(\mu, \tau) \geq 0$, $(\mu, \tau) \neq (0,0)$;
4. *MON*1: $d_{u_n, S, S^*} < 3$;
   *MON*2: $S^* = 0 \Rightarrow d_{u_n, S} < 2$;
   *MON*3: $S = 0 \Rightarrow d_{u_n, S^*} < 2$;
   *MON*4: $d_{u_n, S} = d_{u_n, S^*} = d_{S, S^*} = d_{u_n, S, S^*} = 2 \Rightarrow \exists p, q$: $p \succcurlyeq q \,\&\, p \succcurlyeq^* q \,\&\, p \succcurlyeq_n q$ *with one inequality strict.*

The proof of this proposition is in Appendix A.

Discussion of the Axiom

The Axiom follows from three different trends of thought:

(α) A first is the idea of "monotonicity," or "positive association:"[31] if an individual changes his preferences between p and q, while the rest of the profile remains unchanged, then society's preferences between p and q do not change in the opposite direction. When further the individual is totally indifferent

[27] Just the first part, together with the very weak $p \sim^* q \,\&\, p \sim q \Rightarrow p \sim_n q$, also implies MON.

[28] And every $(\preccurlyeq, \preccurlyeq^*, \preccurlyeq_n)$ satisfying MON satifies the conditions of either fn. 26 or fn. 27.

[29] I.e., computations are done modulo additive constants.

[30] These are analytic conditions, to facilitate understanding and working with the axiom. Condition 1 is the most symmetric; among the more explicit conditions, 2 is the simplest and preserves the symmetry between S and S^*, 3 goes in the direction of natural causality, from S^* and u_n to S, and 4 gives the completely explicit decomposition of the axiom into four mutually disjoint geometric configurations of (S, S^*, u_n) to be excluded.

[31] In Dhillon and Mertens (1997) the correct interpretation of "Positive Association" in a framework of preferences over lotteries is discussed, and it is shown that only the present formulation resists contradiction with Pareto.

before the change, this is equivalent to our above Pareto-like axiom. But this is merely a nonnegative association, being quite compatible with a constant SWF, completely unaffected by any changes in the profile. The minimal expression of a truly positive association is our goodwill axiom. This together with consistency yields (almost) the full strong form of monotonicity (translating as $S = \mu u_n + \tau S^*$ with $\mu > 0, \tau > 0$).

(β) One could also require that society's preferences between p and q be unaffected if the individual is indifferent between them before and after the change. That would be a very particular case of Arrow's IIA. This postulates indeed that social preferences between p and q are determined solely by the profile of individual preferences between them. Its basic difficulty is that, without the preferences themselves between p and q changing, their *intensity* might very well have changed, this being reflected in the SWF. E.g., even in Arrow's framework, in one case p might be at the top of someone's list and q at the bottom, while in the other case both would be at the bottom, in the same order. In our framework of preferences over lotteries, such differences become even more meaningful, identifying the top of the list with all "uncertain prospects" preferred to a lottery where he is a dictator with probability $1 - \varepsilon$ and otherwise is sent for a life-term of forced labor in Siberia, and the bottom with those less preferred than the reverse lottery.

This difficulty disappears in only one very particular case where the two profiles $(\preccurlyeq_n)_{n \in N}$ and $(\preccurlyeq_n^*)_{n \in N}$ that coincide on p and q have the further property that, for every individual n, either $\preccurlyeq_n = \preccurlyeq_n^*$, or both $p \sim_n q$ and $p \sim_n^* q$. More precisely, in the framework of the MON axiom, this means that

$$[p \sim_n q \,\&\, p \sim_n^* q] \Rightarrow [p \succ q \Rightarrow p \succ^* q].$$

Indeed, for player n, intensities are zero anyway, and the others did not change. So this axiom is the only—extremely particular—case of IIA that is not vulnerable to this criticism based on intensities. Observe that its particular case where the individual is totally indifferent before the change is our consistency axiom. Since this translates (Farkas) into $S = \mu u_n + \tau S^*$ with $\tau > 0$, we also get, for some other u_n', $S' = \mu' u_n' + \tau' S^*$ with $\tau' > 0$: eliminating S^* between both equations yields a relation implying the general requirement. So the consistency axiom is the exact expression of this "only nonvulnerable aspect of IIA."

Recall that, as observed before, MON already follows from either goodwill, or consistency, or the Pareto-like axiom—just on their own.

(γ) The "extended Pareto" axiom requires that, when partitioning society into subgroups, the social preferences satisfy Pareto as a function of those of the subgroups: it is really the logical expression of what is meant by aggregating preferences.[32] On the face of it, it relates the SWF's of different societies. But when dealing with a single society and a single SWF, one can identify the social preferences of a subgroup with those the full society would have if all nonmembers of the subgroup were totally indifferent. With this interpretation, it be-

[32] Yielding thus I. Kant's ideal of a categorical imperative, despite Harsanyi (1976, Ch. III).

comes a (multi-profile) axiom for a single-society and a single SWF.[33] Particularized to the case where one of the subgroups is a single individual, it becomes equivalent to the monotonicity axiom, in its full strong form above if one takes strict Pareto, and if one just assumes weak Pareto it still implies our above Pareto-like axiom.

It is to allow also for this "extended Pareto" interpretation of the axiom that we had to restrict all comparisons between two different preferences of an individual to only those cases where one of the two equals total indifference—all other interpretations were perfectly compatible with general comparisons. "Extended Pareto" is further exploited in a similar context in Dhillon (1998).

3. BASIC IDEAS UNDERLYING THE PROOF

As a lighthouse to the reader when drowning in the proof, we present here a heuristic version in a very simplified setting. It is convenient to start from Pareto. However this is not one of our axioms, so we first describe "why" it follows from the others (monotonicity). We will take for this purpose the following very strong definition of monotonicity (cf. Proposition 1): for any profile $\vec{u}$, denote by $S_{\vec{u}}$ a utility representation of the corresponding social utility, and by $\vec{u}_{-n}$ the profile where n's utility has been set to zero. Then $S_{\vec{u}}$ is of the form $\lambda u_n + \mu S_{\vec{u}_{-n}}$, with $\lambda > 0$, $\mu > 0$. By INDIV, $S_{\vec{0}} = 0$, hence, induction on the number of individuals with nonzero utilities yields $S_{\vec{u}} = \Sigma \lambda_n u_n$ with $\lambda_n > 0\ \forall n$, i.e., strict Pareto.

To explain how monotonicity implies the required form of separability, we will assume a framework that makes the meaning of this separability obvious: we consider only profiles $\vec{u}$ for which the image in utility space of the set of lotteries is full-dimensional. Then the vector $(\lambda_n)_{n \in N}$ is uniquely defined (up to positive multiples), so is a well-defined function of the profile: explicitly, $S_{\vec{u}} = \Sigma_n \lambda_n(u_n, \vec{u}_{-n}) u_n$. The separability property we now want to obtain is that λ_n depends only on u_n, not on $\vec{u}_{-n}$.

Observe that monotonicity implies that $\Sigma_n \lambda_n(u_n, \vec{u}_{-n}) u_n = S_{\vec{u}} = \lambda u_1 + \mu S_{\vec{u}_{-1}} = \lambda u_1 + \mu \Sigma_{n \neq 1} \lambda_n(u_n, \vec{u}_{-(n,1)}) u_n$. By the uniqueness of the coefficients,

$$\frac{\lambda_2\big(u_2, \vec{u}_{-2}\big)}{\lambda_3\big(u_3, \vec{u}_{-3}\big)} = \frac{\lambda_2\big(u_2, \vec{u}_{-(1,2)}\big)}{\lambda_3\big(u_3, \vec{u}_{-(1,3)}\big)}$$

is independent of u_1, hence of u_n for all $n \notin \{2,3\}$:

$$\frac{\lambda_i\big(u_i, \vec{u}_{-i}\big)}{\lambda_j\big(u_j, \vec{u}_{-j}\big)} = \frac{\lambda_i\big(u_i(u_j, 0, \dots)\big)}{\lambda_j\big(u_j, (u_i, 0, \dots)\big)}.$$

[33] RU also satifies the extended monotonicity property, where the social preferences of the partition elements play the role of the individual preferences in the strong form of MON. But this extended monotonicity axiom (with INDIV) is equivalent to the more transparent extended Pareto axiom: since it suffices for each one to require it for a partition of society into two subgroups, the equivalence reduces to that of Pareto and the full strong form of MON when $\# N = 2$.

It should thus be clear that λ_i can only depend on u_i. Formally, let

$$F_i(u_i) = \frac{\lambda_i(u_i,(u_1,0,\dots))}{\lambda_1(u_1,(u_i,0,\dots))},$$

for some fixed $u_1 \neq 0$. Dividing the numerator and denominator of $\lambda_i(u_i,\vec{u}_{-i})/\lambda_j(u_j,\vec{u}_{-j})$ by $\lambda_1(u_1,\vec{u}_{-1})$ we get that:

$$\frac{\lambda_i\left(u_i,\vec{u}_{-i}\right)}{\lambda_j\left(u_j,\vec{u}_{-j}\right)} = \frac{F_i(u_i)}{F_j(u_j)},$$

i.e., $\lambda_i(u_i,\vec{u}_{-i})/F_i(u_i)$ is independent of i. Call it G; then $\lambda_i(u_i,\vec{u}_{-i}) = F_i(u_i)G(\vec{u})$. Hence $S_{\vec{u}}$ is represented by $\Sigma_n F_n(u_n)u_n$.

Thus, monotonicity implies first Pareto, and then the required separability of the coefficients.

Let us now look at the implications of IRA (plus Pareto). Henceforth we will use 0–1-normalized versions of individual utilities. Fix the range R (in utility space) of some profile—and recall it is assumed full-dimensional. Assume it has less extreme points than there are alternatives. So for any profile $\vec{u}$ with this range, there exists some alternative a_0 such that $\vec{u}_{a_0}$ is in the convex hull of the $(\vec{u}_a)_{a \neq a_0}$. So a_0 is unanimously indifferent to some lottery on the other alternatives. When $\vec{u}_{a_0}$ varies throughout R this remains true, and individual preferences on $\Delta^{C(a_0)}$ remain constant. By IRA, the resulting social preferences on $\Delta^{C(a_0)}$ are therefore independent of this lottery. Since the image in utility space of $\Delta^{C(a_0)}$ equals the full-dimensional set R, those social preferences can by Pareto be identified with a linear functional $\langle \lambda, \cdot \rangle$ on R, which is independent of $\vec{u}_{a_0}$. Further, by Pareto, the social preferences on Δ^A must be represented by the same linear functional, which is unique by the full-dimensionality of R. So the vector $(\lambda_n(\vec{u}))_{n \in N}$ is independent of the choice of $\vec{u}_{a_0}$ in R. It is clear that, by a sequence of such changes, we can transform the profile $\vec{u}$ to any other profile $\vec{v}$ with the same range R. Therefore $(\lambda_n(\vec{u}))_{n \in N}$ depends only (up to proportionality) on the range R of $\vec{u}$.

Now we look at what this implies in conjunction with the separability we got from monotonicity. We will use the previous conclusion for profiles where $u_n \equiv 0\ \forall n \notin \{1,2\}$ (so the "full-dimensionality" is to be interpreted in the utility space of 1 and 2).

Fix a utility u_1, and alternatives a_0 and a_1 with $u_1(a_0) = 0$, $u_1(a_1) = 1$. Fix two other alternatives a_2 and a_3 and assume, e.g., $0 < u_1(a_2) < 1$. Let $u_2(a) = u_1(a)$ for $a \neq a_3$, and say $u_2(a_3) > u_1(a_3)$. So the range R of (u_1,u_2) is a triangle, hence has less extreme points than $\#A$. Since $0 < u_1(a_2) < 1$, we can move $u_1(a_2)$ slightly downwards without changing R: $\lambda_1(u_1 - \varepsilon\delta_{a_2})/\lambda(u_2)$ is independent of ε, hence so is $\lambda_1(u_1 - \varepsilon\delta_{a_2})$. Applying this argument for appropriate choices of u_2, a_3, and of $u_2(a_3)$ we conclude that $u_1(a_2)$ can be perturbed in each direction without affecting $\lambda_1(u_1)$. So it can be perturbed to any other value in $]0,1[$, and thus by continuity, to any other value, without affecting λ_1. By finitely many such changes, we can transform u_1 to any other utility: λ_1 is constant—and so is any other λ_n.

Anonymity yields then directly that the λ's are independent of n, hence RU.

4. THE THEOREM

THEOREM 1: *RU satisfies the axioms. If $N \geq 3$ and $A \geq 5$, any SWF satisfying the axioms coincides with RU whenever some two individuals do not have equal or opposite preferences.*

PROOF: The first sentence is obvious. For the reverse, let henceforth *individual* utilities be 0–1-normalized or identically zero. We will need the following definition:

DEFINITION 2: For any profile $\vec{u}$, $N_0 = \{n | u_n = 0\}$, $N_+ = CN_0$, $\nu_{\vec{u}} = \#N_+$, $U = \sum_n u_n$, and $S_{\vec{u}}$ (S in short) stands for the corresponding social preferences. $\vec{u}$ is exceptional if $\#A = 4$, $d_{\vec{u}} = 2 < \nu_{\vec{u}}$, U has a single maximum and a single minimum, say a_1 and a_2, and either $u_n \neq 0 \Rightarrow 0 < u_n(a_i) < 1$ for $i = 1, 2$, or all but one of the nonindifferent individuals share a single preference or its opposite, under which neither a_1 nor a_2 is a most or least preferred alternative.

The result will now follow from the following proposition, of which the proof is given in Appendix B:

PROPOSITION 2: *RU holds for all profiles with $d_{\vec{u}} > 2$. More specifically, if $\#N \geq 3$ and $\#A \geq 4$:*

(α) *either $d_{\vec{u}} \neq 1 \& \vec{u}$ nonexceptional or $\nu_{\vec{u}} < \#N$ implies $S_{\vec{u}} = U$;*

(β) *$S_{\vec{u}} = \sum_n \lambda_n u_n$;*

(γ) *$\vec{u}$ nonexceptional $\Rightarrow S_{\vec{u}} \in \{0, U\}$;*

(δ) *$\vec{u}$ exceptional $\Rightarrow S_{\vec{u}} \notin \{0, -U\}$.*

Indeed there are no exceptional profiles for $\#A > 4$, and our condition means $d_{\vec{u}} > 1$ or $0 < \nu_{\vec{u}} < \#N$. *Q.E.D.*

5. NECESSITY OF THE AXIOMS

Below we give for each axiom one or more examples showing it to be necessary, even when the other axioms are taken in their strongest form (and, if relevant, even when assuming in addition strict Pareto).

INDIV: $S = V$, or $V + U$, where V_a is "imposed" in Arrow's words, an ethical norm.

NONT: $S = 0$.

NOILL: $S = -U$ if $v_{\vec{u}} < k$, $= U$ otherwise.

ANON: $S = \sum_n \lambda_n u_n$ with $\lambda_n > 0$.[34]

CONT: $S = \sum_v \alpha_{n(v)} v$ with $\alpha_{n+1} > \alpha_n, \alpha_0 = 0$ and $n(v) = \#\{n | u_n = v\}$. For $\alpha_n = n^p$ the SWF is in addition invariant under replication. For $p \to \infty$ we get S determined by "plurality voting."

IRA: $S = \sum_n u_n / q(u_n)$, with $q(u_n) = \sqrt{\sum_a w(a)[u_n(a) - \mu_n]^2}$ (if nonzero), $\mu_n = \sum_a w(a) u_n(a)$, and $w(a) > 0, \sum_a w(a) = 1$. With w constant neutrality is also satisfied, but this needs finiteness of A, and is sensitive to duplications of alternatives. Note the asymmetric treatment of pure and random prospects, and the uncountable pathology for A when u_n is only w-a.e. constant.

MON: $S = \sum_n \lambda_n u_n$ with λ the gradient, at the maximizer in the closure of $\bar{R}_{\vec{u}}$ of $R_{\vec{u}}$, of the Nash product for the nonindifferent players, with $[\inf_a u_n(a)]_{n \in N}$ as status quo.

This example is typical of an otherwise well-behaved SWF that fails MON; but it fails MON1, MON2, and MON3. Cf. Appendix C for examples showing that each of MON1-MON4 is necessary.

6. THE EXCEPTIONS

For the "exceptional profiles" in dimension 2, they form such a small subset, and only when there are exactly 4 alternatives, that strengthening the axiom system just to cover it is not worthwhile. Indeed it disappears just by invoking either a dummy axiom (adding a totally indifferent individual does not affect social preferences), or a "dummy-alternative" axiom (duplicating an alternative does not affect social preferences). But this would take us out of the framework of fixed sets of individuals and of alternatives; if one really wants to get rid of it within the present framework, only CONT can be usefully strengthened: define $\preccurlyeq^i$ to *converge separately* to a nonconstant $\preccurlyeq^\infty$ if there exist utility representations and $a_0 \in A$ such that all representations coincide on $C\{a_0\}$, and they converge at a_0. Strengthen then CONT by just requiring social preferences to converge separately. This suffices.

[34] We feel confident that the rest of the axiom system is reduced to being roughly both minimal and optimal. But for ANON no such effort was made. As said at the end of the introduction, we expect the result to remain true without ANON, when individual weights multiply the normalized utilities. A first advantage would be that the full axiom system becomes then strictly a weakening of Arrow's. Another advantage is that it no longer depends on our argument in the introduction that only individual preferences should occur as primitives; the weights serve then to represent the individuals' relative efficiencies as pleasure machines, as determined, e.g., from their expressive reactions. But as a side-benefit RU would already follow with a weaker, single-profile version of ANON: that a symmetric profile has a symmetric solution—formally, e.g., let $f\colon A \to A$ be its own inverse, and assume f permutes u_{n_0} with u_{n_1}, while leaving the rest of $\vec{u}$ unchanged. Then $a \sim f(a)\ \forall a$.

For the profiles of dimension 1,[35,36] clearly invoking a dummy axiom would solve this case too. Without that, more continuity wouldn't help here; but the extremely mild additional axiom, "if when n is totally indifferent society is not so; then if n switches exactly to those preferences of society the latter does not become totally indifferent," excludes already all those exceptions save for the case of an odd number of individuals, where nobody abstains, and the size of the majority is exactly 1. To get rid of this last case as well would need a strong axiom: that if society is totally indifferent when individual 1 is, then when 1 is no longer so, society also is not.[37] This is a strong axiom in the sense that it is the only one in our system that requires a non-null effect of a single individual's preferences on society's.[38] And indeed such votes, when they occur, do tend to raise eyebrows.

Dept. of Economics, University of Warwick, Coventry CV4 7AL, U.K.; ecsbk@csv.warwick.ac.uk

and

CORE, Université Catholique de Louvain, 34, Voie du Roman-Pays, B-1348 Louvain-la-Neuve, Belgique; jfm@core.ucl.ac.be

Manuscript received June, 1994; final revision received June, 1998.

APPENDIX A: PROOF OF PROPOSITION 1

We prove $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow \text{MON} \Rightarrow 4 \Rightarrow 1$.

$1 \Rightarrow 2$: Either $\mu = 0$, so $\alpha \neq 0, \tau \neq 0$, and $\alpha S + \tau S^* = 0$ (i.e. $S = \pm S^*$), or (normalizing) $\mu = -1$; then we get $u_n = \alpha S + \tau S^*$ with $\max(\alpha, \tau) \geq 0$.

$2 \Rightarrow 3$: If $S = \pm S^*$ set $\mu = 0$, $\tau = \pm 1$. Otherwise $u_n = \alpha S + \tau S^*$, $\max(\alpha, \tau) \geq 0$—and $(\alpha, \tau) \neq (0,0)$ because $u_n = 0$ implies by definition $S = S^*$. Then if $\alpha = 0$ we have $\tau \neq 0$, so $S^* = \pm u_n$, and otherwise $S = (1/\alpha)u_n - (\tau/\alpha)S^*$: since $(1/\alpha) \neq 0$ we only need that $\max(1/\alpha, -\tau/\alpha) \geq 0$, which follows from $\max(\alpha, \tau) \geq 0$.

$3 \Rightarrow \text{MON}$: If $S^* = \pm u_n$, MON holds, the premises $p \succ_n r, p \sim^* r$ being incompatible. Else $S = \mu u_n + \tau S^*$, $(\mu, \tau) \neq (0,0)$, $\max(\mu, \tau) \geq 0$. Then, in case of strict preferences, $r \succ^* q$, $q \succ_n r$, and $r \sim q$ implies μ and τ have the same sign. This can neither be zero because $(\mu, \tau) \neq (0,0)$, nor negative because $\max(\mu, \tau) \geq 0$. Thus $\mu > 0$. So trivially $p \sim^* r$ & $p \succ_n r \Rightarrow p \succ r$. And in case $q \sim^* r$, $q \succ_n^* r$ and $q \sim r$ yields $\mu = 0$, hence $\tau \neq 0$, i.e., $S = \pm S^*$ so that $p \sim^* q \Rightarrow p \sim q$.

[35] We have a bit more here from our proofs: as in Part I, $S = U$ or $S = 0$ depending only on the range of U.

[36] Observe these are essentially the profiles covered by May's theorem, since there exist two alternatives (certainly when $\#A$ is finite) such that any other alternative is unanimously equivalent to one of them.

[37] With this strengthening, our monotonicity axiom becomes similar (in a 2 alternative world) to May's "positive responsiveness," or in general, to a weak form of Positive Association. However, as seen before, we do not assume neutrality.

[38] Substantial effort was put into the axiom system to achieve this: except for NONT, which is completely minimal, the rest of the axiom system is fully compatible even with a constant SWF. The importance of this aspect should be obvious in a world where majority rules are so frequent—under which typically a single individual out of millions can change his preferences arbitrarily without affecting social preferences.

MON $\Rightarrow$ 4: MON1: If there were no linear relation, $C=\{(\langle u_n,p\rangle,\langle S^*,p\rangle,\langle S,p\rangle)|p\in\Delta^A\}$ would be a full-dimensional convex set in $\mathbb{R}^3$. So it contains a small cube, within which we can find three points, say x_p, x_q, and x_r, with $x_p^1=x_q^1>x_r^1$, $x_r^2=x_p^2>x_q^2$, $x_p^3<x_q^3=x_r^3$, contradicting MON.

MON2: Assume the (convex) range $C'=\{(\langle u_n,p\rangle,\langle S,p\rangle)|p\in\Delta^A\}$ is full-dimensional, and $S^*=0$. C' contains then 3 points x_p, x_q, and x_r, with $x_p^1=x_q^1>x_r^1$, $x_p^2<x_q^2=x_r^2$, contradicting MON.

MON3: If $C''=\{(\langle u_n,p\rangle,\langle S^*,p\rangle)|p\in\Delta^A\}$ is full-dimensional, it contains 3 points with $x_p^1=x_q^1>x_r^1$, $x_q^2<x_p^2=x_r^2$, so $S\neq 0$ by MON since $p\succ q$ or $q\nsucc r$.

MON4: If MON4 is false, the dimension coordinates yield a unique linear relation between u_n, S, and S^*, with all coefficients nonzero—and of the same sign since if one of the utility functions was, after scaling, the sum of the two others, there would be a pair (p,q) with $p\succ q$ for each of the latter (dimensionality), hence also for their sum. So $S=\mu u_n+\tau S^*$ with $\mu<0,\tau<0$. Since $d_{u_n,S^*}=2$, take $p\sim_n q\succ_n r$ and $p\sim^* r\succ^* q$. By MON, we need $r\nsim q$. So, either $r\succ q\succ p$ (and then for an appropriate convex combination $\bar{r}$ of r and p we have $\bar{r}\sim q$, and hence a contradiction with $\bar{r}$ instead of r) or we have $q\succ r\succ p$ (then for a convex combination $\bar{q}$ of q and p we have $r\sim\bar{q}$, and hence a contradiction with $\bar{q}$ instead of q).

$4\Rightarrow 1$: By MON1, $\exists(\alpha,\mu,\tau)\neq(0,0,0)$ with $\alpha S+\mu u_n+\tau S^*=0$. If exactly one of the three coefficients is nonzero, then 1 is satisfied: indeed, the corresponding utility would be identically zero, so, if $u_n=0$, we would have $S=S^*$ by definition, hence another solution $(1,1,-1)$ satisfying 1. If $S^*=0$, we would have by MON2 that either $u_n=0$, so we are in the previous case, or $S=\lambda u_n$, hence another solution $(1,-\lambda,-1)$ satisfying 1; similarly if $S=0$, using MON3 this time.

Further, 1 is trivially satisfied if exactly two of the coefficients are nonzero.

So assume all three coefficients are nonzero, for any such relation, and all of the same sign, lest 1 still holds. Hence the relation, $\alpha S+\mu u_n+\tau S^*=0$ with $\alpha>0$, $\mu>0$, $\tau>0$, is unique; else a linear combination would have a zero coefficient. Thus $d_{u_n,S,S^*}=2$. Since any variable can be expressed linearly in terms of the other two, any pair of variables has dimension two, lest $d_{u_n,S,S^*}\leq 1$.

So the pair (p,q) asserted in MON4 contradicts that $\alpha>0$, $\mu>0$, $\tau>0$.

APPENDIX B: Proof of Proposition 2

Notation

$\wp^*$ is the set of nonconstant preferences. $S_u, S_{u,v}, S_{u,v,w}$ are used for $S_{\bar{\pi}}$ when everyone is totally indifferent but for one individual having utility u, or two having u and v, etc. By ANON, names of individuals don't matter, hence justifying the notation. For $A'\subset A$, $\pi\colon \complement A'\to\Delta^{A'}$, and u normalized on A', let $u^\pi=u$ on A', else $u_a^\pi=\sum_{\alpha\in A'}\pi_\alpha^a u_\alpha$.

Lemma 1: *$S_{u,v}=u+v$ except possibly when $v=\pm u\neq 0$.*

The proof consists of five parts.

Part I

(α) *$\exists\lambda(R)\in\{-1,0,1\}$ for $]0,1[\subseteq R\subseteq[0,1]$, such that $S_u=\lambda(R_u)u$ $\forall u$.*

(β) *If $R'\subseteq R$, then $\lambda(R)\in\{0,\lambda(R')\}$. So, either $\lambda(R)\geq 0$ $\forall R$, or $\lambda(R)\leq 0$ $\forall R$.*

(γ) *λ is not identically zero: the two alternatives above are mutually exclusive.*

Proof: INDIV takes care of $u=0$. Hence $S_u=\lambda_u u$ by MON2, where λ_u has values in $\{-1,0,1\}$. For v on $A'\subset A$, S_{v^π} and $S_{v^{\pi'}}$ coincide by IRA on $\Delta^{A'}$, so $\lambda_{v^\pi}=\lambda_{v^{\pi'}}$: if the ranges of u and $\bar{u}$ are already obtained on $\{a|u_a=\bar{u}_a\}$, then $\lambda_u=\lambda_{\bar{u}}$.

Thus λ_u depends only on the range of $u([0,1],]0,1], [0,1[$ or $]0,1[)$: If $[0,1]$, then by the above $\lambda_u = \lambda_{\tilde{u}}$ for some indicator function $\tilde{u}$ of a singleton, and any two such $\tilde{u}$ have the same λ as their sum w.

If it is $[0,1[$, there exists a_i with $u(a_i)$ strictly increasing to 1, and $u(a_0) = 0$. By the above make $u(a) = 0$, elsewhere preserving λ. Similarly for v with a sequence b_i. Make further a_i and b_i disjoint (if they have a common subsequence, take its even terms for a_i and the others for b_i, else delete the common terms). Again u and v have the same λ as their sum w. The case of $]0,1]$ is dual.

For $]0,1[$, use four mutually disjoint sequences a_i, $\underline{a}_i$, b_i and $\underline{b}_i$, such that $u(a_i)$ and $v(b_i)$ increase to 1, $u(\underline{a}_i)$ and $v(\underline{b}_i)$ decrease to 0, $u(a) = v(a) = 1/2$ elsewhere; set $w = u$ at a_i and $\underline{a}_i$, $= v$ elsewhere.

Hence α. CONT implies now β, since some u with range R can be specially approximated by $\tilde{u}$'s with range R', if any. As to γ: otherwise $S_u = 0\ \forall u$, hence inductively $S_{u_1,\dots,u_n} = 0\ \forall n$ (thus contradicting NONT). Indeed, assume this; then $S_{u_1,\dots,u_{n+1}} = \mu u_{n+1}$ by MON2, so $= \mu_i u_i \forall i$; thus $= 0$ if $d_{u_1,\dots,u_{n+1}} > 1$. The case $d_{\bar{u}} = 1$ follows now by CONT, hence the induction. *Q.E.D.*

PART II: *Implications of MON: the product relation.*

Step A

(α) *If* $d_{u,v} = 2$, *there exist* $\lambda_{u,v}, \lambda_{v,u}$ *(unique) such that* $S_{u,v} = \lambda_{u,v}u + \lambda_{v,u}v$.
(β) $(\lambda_{u,v}, \lambda_{v,u}) = (0,0) \Rightarrow (\lambda(R_u), \lambda(R_v)) = (0,0)$, $\lambda(R_u) = 0 \Rightarrow \lambda_{u,v} = 0\ \forall v$.
(γ) *If* $\lambda(R) \geq 0\ \forall R$, *then* $\max(\lambda_{u,v}, \lambda_{v,u}) \geq 0$.

PROOF: By Proposition 1.3, and part I, $S_{u,v} = \mu u + \tau\lambda(R_u)v$ (since $d_{u,v} = 2$), with $(\mu, \tau) \neq (0,0)$, and $\max(\mu, \tau) \geq 0$. Hence α. In particular $\lambda_{v,u} = \tau\lambda(R_v)$, thus the second part of β. If $\lambda_{u,v} = \lambda_{v,u} = 0$, then $(\mu, \tau) \neq (0,0)$ implies $\lambda(R_v) = 0$. Duality $\lambda(R_u) = 0$, hence β. γ is immediate. *Q.E.D.*

Step B

$\lambda_{u,v}\lambda_{v,w}\lambda_{w,u} = \lambda_{u,w}\lambda_{w,v}\lambda_{v,u}$ *whenever* $d_{u,v,w} = 3$.

PROOF: Consider ($\#N \geq 3$) the profile (u, v, w): by A and Proposition 1.3 we obtain

$$S_{u,v,w} = \mu_1 u + \tau_1(\lambda_{v,w}v + \lambda_{w,v}w) \tag{1}$$

$$= \mu_2 v + \tau_2(\lambda_{u,w}u + \lambda_{w,u}w) \tag{2}$$

$$= \mu_3 w + \tau_3(\lambda_{u,v}u + \lambda_{v,u}v) = \lambda_1 u + \lambda_2 v + \lambda_3 w. \tag{3}$$

Since $d_{u,v,w} = 3$ the λ_i are unique. If two of them, say λ_2 and λ_3, were zero, $(\mu, \tau) \neq (0,0)$ would imply in (2) and (3) $\lambda_{w,u} = \lambda_{v,u} = 0$, hence the equality. Else $\tau_i \neq 0\ \forall i$, so $\lambda_{v,w} = \lambda_2/\tau_1$, $\lambda_{w,v} = \lambda_3/\tau_1$, $\lambda_{u,w} = \lambda_1/\tau_2$, $\lambda_{w,u} = \lambda_3/\tau_2$, $\lambda_{u,v} = \lambda_1/\tau_3$, $\lambda_{v,u} = \lambda_2/\tau_3$, thus both members equal $\lambda_1\lambda_2\lambda_3/\tau_1\tau_2\tau_3$. *Q.E.D.*

Step C

$\lambda_{u,v}\lambda_{v,w}\lambda_{w,u} = \lambda_{u,w}\lambda_{w,v}\lambda_{v,u}$ whenever $d_{u,v} = d_{v,w} = d_{w,u} = 2$.

PROOF: Assume $d_{u,v,w} = 2$, using B. Then ($\#A \geq 4$) there exists x with $d_{u,v,w,x} = 3$ (so $d_{u,v,x} = d_{v,w,x} = d_{w,u,x} = 3$). If the equality does not hold, we get by B:

$$\lambda_{u,v}\lambda_{v,x}\lambda_{x,u} = \lambda_{u,x}\lambda_{x,v}\lambda_{v,u}, \tag{4}$$

$$\lambda_{v,w}\lambda_{w,x}\lambda_{x,v} = \lambda_{v,x}\lambda_{x,w}\lambda_{w,v}, \tag{5}$$

$$\lambda_{w,u}\lambda_{u,x}\lambda_{x,w} = \lambda_{w,x}\lambda_{x,u}\lambda_{u,w}. \tag{6}$$

By multiplying term-wise, and subtracting,

$$[\lambda_{u,v}\lambda_{v,w}\lambda_{w,u} - \lambda_{v,u}\lambda_{w,v}\lambda_{u,w}](\lambda_{u,x}\lambda_{v,x}\lambda_{w,x}\lambda_{x,u}\lambda_{x,v}\lambda_{x,w}) = 0.$$

Since by assumption the first bracket is not zero, we conclude that, whenever $d_{u,v,w,x} = 3$ one has

$$(\lambda_{u,x}\lambda_{v,x}\lambda_{w,x})(\lambda_{x,u}\lambda_{x,v}\lambda_{x,w}) = 0.$$

Since the equality doesn't hold, A.β yields $\lambda(R_u) \neq 0$, $\lambda(R_v) \neq 0$, $\lambda(R_w) \neq 0$: by A.β, $\lambda_{s,t} = 0 \Rightarrow \lambda_{t,s} \neq 0$ for all above pairs (s,t). Then, e.g., $\lambda_{u,x} = 0$ implies $\lambda_{x,u} \neq 0$, so $\lambda_{u,v}\lambda_{v,x} = 0$ by (4) and $\lambda_{u,w}\lambda_{w,x} = 0$ by (6): Either $\lambda_{u,v} = \lambda_{u,w} = 0$, thus equality; or $\lambda_{u,v} = \lambda_{w,x} = 0$, so $\lambda_{x,w} \neq 0$ and (5) imply $\lambda_{v,x}\lambda_{w,v} = 0$, and since $\lambda_{u,v} = \lambda_{w,v} = 0$ yields equality we get $\lambda_{v,x} = \lambda_{w,x} = 0$; or $\lambda_{v,x} = \lambda_{u,w} = 0$, where we conclude similarly that $\lambda_{v,x} = \lambda_{w,x} = 0$; or finally again $\lambda_{v,x} = \lambda_{w,x} = 0$: $\lambda_{u,x} = 0$ implies $\lambda_{u,x} = \lambda_{v,x} = \lambda_{w,x} = 0$. The same clearly applies if $\lambda_{v,x} = 0$ or $\lambda_{w,x} = 0$, and similarly $\lambda_{x,u} = 0$ yields $\lambda_{x,v} = \lambda_{x,w} = 0$.

So $d_{u,v,w,x} = 3$ implies either $\lambda_{u,x} = \lambda_{v,x} = \lambda_{w,x} = 0$ or $\lambda_{x,u} = \lambda_{x,v} = \lambda_{x,w} = 0$.

Let w^ε converge specially to w, with $d_{u,v,w,w^\varepsilon} = 3$ (possible by $\#A \geq 4$). If (along a subsequence) $\lambda_{u,w^\varepsilon} = \lambda_{v,w^\varepsilon} = 0$, then $\lambda_{w^\varepsilon,u} \neq 0$, so $S_{u,w^\varepsilon} = \lambda_{u,w^\varepsilon}u + \lambda_{w^\varepsilon,u}w^\varepsilon = \pm w^\varepsilon$, and along a subsequence with constant sign $S_{u,w} \in \{w, 0, -w\}$ by CONT, $\lambda_{u,w} = 0$. Similarly $\lambda_{v,w} = 0$, thus the claim. Else $\lambda_{u,w^\varepsilon} \neq 0 = \lambda_{w^\varepsilon,u}$, so $S_{u,w^\varepsilon} = \pm u$, hence as above $\lambda_{w,u} = 0$, similarly; again equality follows. *Q.E.D.*

Step D

There exist $\lambda_{u,u} > 0$ and $\lambda_{u,-u}$ such that:
(α) $u \neq 0, v \neq 0, w \neq 0 \Rightarrow \lambda_{u,v}\lambda_{v,w}\lambda_{w,u} = \lambda_{u,w}\lambda_{w,v}\lambda_{v,u}$;
(β) $\lambda(R_u) = 0 \Rightarrow \lambda_{u,v} = 0\ \forall v \neq u$;
(γ) $(\lambda_{u,v}, \lambda_{v,u}) = (0,0) \Rightarrow (\lambda(R_u), \lambda(R_v)) = (0,0)$;
(δ) *if* $\lambda_{-u,v}\lambda_{v,u} = 0\ \forall v \neq \pm u$ *and* $\lambda(R_u) \neq 0$, *then* $\lambda_{u,-u} = 1$.

Proof: If two of the utilities u, v, and w are equal, then α holds. And since $\lambda_{u,u} > 0$ is compatible with β to δ, we can henceforth assume $u \neq v, v \neq w, w \neq u$. Since all relations hold, by A and C, if $\lambda_{u,-u}$ is not involved, we can for instance assume $w = -u$: we need that $\forall v \notin \{u, -u\}$

$$\lambda_{u,-u}\lambda_{-u,v}\lambda_{v,u} = \lambda_{-u,u}\lambda_{u,v}\lambda_{v,-u} \tag{7}$$

and that β and γ hold with $v = -u$, and we need δ. If $\lambda(R_u) = \lambda(R_{-u}) = 0$, set $\lambda_{u,-u} = \lambda_{-u,u} = 0$; if $\lambda(R_u) = 0 \neq \lambda(R_{-u})$, set $\lambda_{u,-u} = 0$, $\lambda_{-u,u} = 1$. Then all conditions are satisfied, because $\lambda(R_u) = 0 \Rightarrow \lambda_{u,v} = 0\ \forall v \neq \pm u$. Thus the case $\lambda(R_u) \neq 0$, $\lambda(R_{-u}) \neq 0$ remains.

δ specifies first $\lambda_{u,-u} = \lambda_{-u,u} = 1$ if $\forall v \neq \pm u\, \lambda_{-u,v}\lambda_{v,u} = \lambda_{u,v}\lambda_{v,-u} = 0$, which is clearly compatible with the other requirements. Otherwise, we have e.g. $\lambda_{u,v}\lambda_{v,-u} = 0\ \forall v \neq \pm u$, $\exists v \neq \pm u$: $\lambda_{-u,v}\lambda_{v,u} \neq 0$. δ specifies then $\lambda_{-u,u} = 1$—and (7) imposes $\lambda_{u,-u} = 0$—which again satisfies all requirements.

In the remaining case $\exists v \neq \pm u$: $\lambda_{u,v}\lambda_{v,-u} \neq 0$ and $\exists v \neq \pm u$: $\lambda_{-u,v}\lambda_{v,u} \neq 0$. We need then a nonzero solution $(\lambda_{u,-u}, \lambda_{-u,u})$ to the system (7), i.e. that its rank be ≤ 1. So for v and $w \neq \pm u$, $v \neq w$, the corresponding determinant must be zero:

$$\lambda_{u,v}\lambda_{v,-u}\lambda_{-u,w}\lambda_{w,u} \stackrel{?}{=} \lambda_{u,w}\lambda_{w,-u}\lambda_{-u,v}\lambda_{v,u}. \tag{8}$$

If $w \neq -v$, $\lambda_{w,u}\lambda_{u,v}\lambda_{v,w} = \lambda_{v,u}\lambda_{u,w}\lambda_{w,v}$ and $\lambda_{v,-u}\lambda_{-u,w}\lambda_{w,v} = \lambda_{w,-u}\lambda_{-u,v}\lambda_{v,w}$, by C. Hence $[\lambda_{u,v}\lambda_{v,-u}\lambda_{-u,w}\lambda_{w,u}](\lambda_{v,w}\lambda_{w,v}) = [\lambda_{u,w}\lambda_{w,-u}\lambda_{-u,v}\lambda_{v,u}](\lambda_{v,w}\lambda_{w,v})$, multiplying term-wise: if (8) did not hold, we must have $\lambda_{v,w}\lambda_{w,v} = 0$. Assume by symmetry that $\lambda_{v,w} = 0$. If also $\lambda_{w,v} = 0$, we obtain $\lambda(R_v) = \lambda(R_w) = 0$ by γ, hence $\lambda_{w,u} = \lambda_{v,u} = 0$ by β, and thus (8). Otherwise, the two equations yield $\lambda_{v,-u}\lambda_{-u,w} = 0 = \lambda_{v,u}\lambda_{u,w}$, and since those factor the left and right-hand members, we get again (8).

Thus to take care of the case $w = -v$, remains, i.e. to show that $d_{u,v} = 2 \overset{?}{\Rightarrow} \lambda_{u,v}\lambda_{v,-u}\lambda_{-u,-v}\lambda_{-v,u} = \lambda_{u,-v}\lambda_{-v,-u}\lambda_{-u,v}\lambda_{v,u}$. Suppose the inequality holds. For w with $w \neq \pm u$, $w \neq \pm v$, we have by C:

$$\lambda_{u,v}\lambda_{v,w}\lambda_{w,u} = \lambda_{u,w}\lambda_{w,v}\lambda_{v,u}, \tag{9}$$

$$\lambda_{-u,w}\lambda_{w,v}\lambda_{v,-u} = \lambda_{-u,v}\lambda_{v,w}\lambda_{w,-u}, \tag{10}$$

$$\lambda_{-u,-v}\lambda_{-v,w}\lambda_{w,-u} = \lambda_{-u,w}\lambda_{w,-v}\lambda_{-v,-u}, \tag{11}$$

$$\lambda_{u,w}\lambda_{w,-v}\lambda_{-v,u} = \lambda_{u,-v}\lambda_{-v,w}\lambda_{w,u}. \tag{12}$$

Multiplying term-wise we get

$$[\lambda_{u,v}\lambda_{v,-u}\lambda_{-u,-v}\lambda_{-v,u} - \lambda_{u,-v}\lambda_{-v,-u}\lambda_{-u,v}\lambda_{v,u}](\lambda_{w,u}\lambda_{u,w})(\lambda_{w,-u}\lambda_{-u,w}) \\ \times(\lambda_{w,v}\lambda_{v,w})(\lambda_{w,-v}\lambda_{-v,w}) = 0.$$

By the inequality, one of the last four brackets must be zero, e.g. $\lambda_{w,u}\lambda_{u,w} = 0$.

Observe that both cannot be zero, otherwise $(\gamma)\,\lambda(R_u) = 0$ would (β) imply our equality. Assume first $\lambda_{u,w} = 0$, so by (9) and (12) $\lambda_{u,v}\lambda_{v,w} = \lambda_{u,-v}\lambda_{-v,w} = 0$. Since $\lambda_{u,v} = \lambda_{u,-v} = 0$ implies equality, either: $\lambda_{u,v} = \lambda_{-v,w} = 0$, hence $\lambda_{w,-v} \neq 0$, so by (11) $\lambda_{-u,w}\lambda_{-v,-u} = 0$, and $\lambda_{-u,w} = 0$ by the inequality; or: $\lambda_{v,w} = \lambda_{u,-v} = 0$, hence $\lambda_{w,v} \neq 0$, so by (10) $\lambda_{-u,w}\lambda_{v,-u} = 0$, and $\lambda_{-u,w} = 0$ by the inequality; or: $\lambda_{v,w} = \lambda_{-v,w} = 0$, thus as above by (10) and (11) $\lambda_{-u,w}\lambda_{v,-u} = \lambda_{-u,w}\lambda_{-v,-u} = 0$ so again $\lambda_{-u,w} = 0$.

Thus, $\lambda_{u,w} = 0$ implies that either $\lambda_{u,v} = \lambda_{-v,w} = 0$ or $\lambda_{v,w} = \lambda_{u,-v} = 0$ or $\lambda_{v,w} = \lambda_{-v,w} = 0$—and also $\lambda_{-u,w} = 0$, hence also the duals of those three alternatives obtained by changing the sign of u. So $\lambda_{v,w} = \lambda_{-v,w} = 0$ follows too; otherwise if e.g. $\lambda_{v,w} \neq 0$ we should have both $\lambda_{u,v} = \lambda_{-v,w} = 0$ and $\lambda_{-u,v} = \lambda_{-v,w} = 0$, hence again equality. So $\lambda_{u,w} = 0 \Rightarrow \lambda_{-u,w} = \lambda_{v,w} = \lambda_{-v,w} = 0$. Similarly $\lambda_{w,u} = 0 \Rightarrow \lambda_{w,-u} = \lambda_{w,v} = \lambda_{w,-v} = 0$.

In conclusion, if our equality does not hold, then for all $w \notin \{u, -u, v, -v\}$: either (a) $\lambda_{u,w} = \lambda_{-u,w} = \lambda_{v,w} = \lambda_{-v,w} = 0$ or (b) $\lambda_{w,u} = \lambda_{w,-u} = \lambda_{w,v} = \lambda_{w,-v} = 0$ with each time *all* λ's mentioned in the other case nonzero (previous argument).

Let then u^ε converge specially to u, so $u^\varepsilon \notin \{u, -u, v, -v\}$, and let $w = u^\varepsilon$; along an appropriate subsequence, we can assume that either (a) holds for all ε, or (b). Assume first (a). Then $S_{u^\varepsilon,v} = \pm u^\varepsilon$, so (along an appropriate subsequence), $S_{u^\varepsilon,v}$ also converges specially, and by CONT we obtain that $S_{u,v} \in \{u, -u, 0\}$, so $\lambda_{v,u} = 0$. Similarly, $\lambda_{-v,u} = 0$, and thus our equality holds. In case (b): $S_{u^\varepsilon,v} = \pm v$, hence $S_{u,v} \in \{v, -v, 0\}$, so $\lambda_{u,v} = 0$. Similarly, $\lambda_{u,-v} = 0$, hence again equality.

Q.E.D.

PART III: *There are maps ξ from $\wp^*$ onto $(X, \geq)$ totally ordered, $F : \wp^* \to \mathbb{R}$, $\varsigma : \wp^* \times \wp^* \to \{-1, 1\}$ such that:*

(α) *if $d_{u,v} = 2$, then $S_{u,v} = \mu_{u,v}u + \mu_{v,u}v$, with $\mu_{u,v} = \varsigma_{u,v}F_u\mathbb{1}_{\xi_u \geq \xi_v}$;*

(β) $\varsigma_{u,v} = \varsigma_{v,u}$;

(γ) $\xi_u < \xi_{-u} \Rightarrow \exists v \neq \pm u : \xi_u \leq \xi_v \leq \xi_{-u}$;

(δ) *if $\lambda(R) = 0$ for some R, then X has a smallest element, denoted $\underline{x}$. Otherwise there is no $\underline{x}$ in X. With this notation, $F_u = 0 \Leftrightarrow \lambda(R_u) = 0 \Leftrightarrow \xi_u = \underline{x}$.*

PROOF: Define the binary relation $\mathscr{R}$ on $\wp^*$ by $u\mathscr{R}v$ iff $(\lambda_{u,v} \neq 0$ or $\lambda_{v,u} = 0)$. By definition, either $u\mathscr{R}v$ or $v\mathscr{R}u$. In particular $u\mathscr{R}u$. Also $u\mathscr{R}v$ and $v\mathscr{R}w$ imply $u\mathscr{R}w$: since $u\mathscr{R}u$, it suffices to prove this when u, v, and w are different. Then, if false, we would have $\lambda_{u,w} = 0$ and $\lambda_{w,u} \neq 0$, hence by D(α), $\lambda_{u,v}\lambda_{v,w} = 0$. So if $\lambda_{u,v} = 0$, $u\mathscr{R}v$ yields $\lambda_{v,u} = 0$, hence $\lambda(R_v) = 0$ by D(γ), thus $\lambda_{v,w} = 0$ by D(β). So anyway $\lambda_{v,w} = 0$. Then $v\mathscr{R}w$ yields $\lambda_{w,v} = 0$. Thus $\lambda(R_w) = 0$ by D(γ), and hence $\lambda_{w,u} = 0$ by D(β), a contradiction.

Thus there is an induced equivalence relation $u \sim v[\equiv u\mathscr{R}v$ and $v\mathscr{R}u]$, and $\mathscr{R}$ induces a total order $\geq$ on the quotient X. ξ denotes the quotient mapping. So

$$\xi_u \geq \xi_v \quad \text{iff} \quad (\lambda_{u,v} \neq 0 \text{ or } \lambda_{v,u} = 0). \tag{13}$$

Consider now some u with $\lambda(R_u)=0$, if any. Then, by D(β) and (13), $\xi_u \leq \xi_v$ for all v: ξ_u is the smallest element $\underline{x}$ of X. Conversely if $\xi_v=\underline{x}$ for some v, then $v \sim u$, meaning that either $\lambda_{u,v} \neq 0$, $\lambda_{v,u} \neq 0$ or $\lambda_{u,v}=\lambda_{v,u}=0$. The first case is impossible by D(β) since $\lambda(R_u)=0$, so by D(γ) we conclude that also $\lambda(R_v)=0$: thus δ is established, except for the equivalence with $F_u=0$. From (1) we obtain:

(14) if $\xi_u > \xi_v$ then $\lambda_{v,u}=0$ and $\lambda_{u,v} \neq 0$,

(15) if $\xi_u = \underline{x}$ then $\lambda_{u,v}=0\ \forall v \neq u$,

(16) if $\xi_u = \xi_v \neq \underline{x}$ then $\lambda_{u,v} \neq 0$.

For each $x \neq \underline{x}$, fix now some $u_0 \in \xi^{-1}(x)$, and let $F_u=\lambda_{u,u_0}/\lambda_{u_0,u}\ \forall u \in \xi^{-1}(x)$. For u and v both in $\xi^{-1}(x)$, define $\delta_{u,v}=\lambda_{u,v}/F_u$. By (16), F is well-defined and nonzero, hence δ is so too. Further D(α) yields, using u_0 for w, that $\delta_{u,v}=\delta_{v,u}$. By this symmetry of δ, we can divide $S_{u,v}$ by $|\delta_{u,v}|$, so $S_{u,v}=\mu_{u,v}u+\mu_{v,u}v$ with $\mu_{u,v}=F_u s_{u,v}$ and $s_{u,v}=\operatorname{sign}(\delta_{u,v})$; i.e., α and β hold when $\xi_u=\xi_v \neq \underline{x}$.

For $u,v \in \xi^{-1}(\underline{x})$, let $F_u=0$ and $s_{u,v}=1$: α and β hold now whenever $\xi_u=\xi_v$, and δ always holds. s is defined on $\cup_x[\xi^{-1}(x)]^2$ and F everywhere. If $\xi_u<\xi_v$, then by (14), α is still true—whatever $s_{u,v}$ is. If $\xi_u>\xi_v$, then $\lambda_{u,v} \neq 0$ by (14), and $F_u \neq 0$ by δ, so α becomes always valid with $s_{u,v}=\operatorname{sign}(\lambda_{u,v}/F_u)$. To preserve β, we need $s_{v,u}$ to be the same—and s is now defined on $\wp^* \times \wp^*$, satisfying β everywhere.

So α, β, and δ are proved. Assume γ false; hence (14) $\forall v \neq \pm u$, $\lambda_{-u,v}\lambda_{v,u}=0$. Also $\lambda(R_u) \neq 0$; otherwise $\xi_v=\xi_u=\underline{x}$ by δ for all v with the same range. And $\lambda_{u,-u}=0$ by (14)—thus contradicting D(δ). *Q.E.D.*

PART IV:: *If $u^\varepsilon \to u$ specially, one of the following holds for a subsequence:*

(α) *ξ_{u^ε} is strictly increasing, $\xi_u=\sup_\varepsilon \xi_{u^\varepsilon}$, and $\xi_v=\xi_u \Rightarrow v=\pm u$;*

(β) *ξ_{u^ε} is strictly decreasing, $\xi_u=\inf_\varepsilon \xi_{u^\varepsilon}$, and $\xi_u=\xi_v \neq \underline{x} \Rightarrow v=\pm u$;*

(γ) *ξ_{u^ε} is strictly decreasing, $\inf_\varepsilon \xi_{u^\varepsilon}=\xi_{-u}>\xi_u$, and $\underline{x}<\xi_v \leq \xi_{-u} \Rightarrow v=-u$;*

(δ) *$\xi_{u^\varepsilon}=\xi_u\ \forall\varepsilon$, and $F_{u^\varepsilon} \to F_u$;*

(ϵ) *$\xi_u \neq \xi_{u^\varepsilon}=x_0 \neq \underline{x}\ \forall\varepsilon$, $]x_0,\xi_u[\,=\,]\xi_u,x_0[\,=\varnothing$, $\xi_v=\xi_u \Rightarrow v=\pm u$, and $F_{u^\varepsilon} \to 0$ if $\xi_u<x_0$, else $\to\infty$;*

(ζ) *$\forall\varepsilon\ \xi_{u^\varepsilon}=x_0>\xi_u=\underline{x}$, and $\xi^{-1}(]\xi_u,x_0[) \subseteq \{-u\}$, and $F_{u^\varepsilon} \to 0$.*

COROLLARY: *If $\exists v \neq \pm u$: $\xi_v=\xi_u \neq \underline{x}$ and $u^\varepsilon \to u$ specially, then $\xi_{u^\varepsilon}=\xi_u$ eventually and $F_{u^\varepsilon} \to F_u$.*

Claim

A sequence in a totally ordered set has a monotone subsequence.

PROOF OF THE CLAIM: Extract a subsequence such that, either $x_i=x_1\ \forall i>1$ (so we have finished), or $x_i>x_1$, or $x_i<x_1$. Repeat with x_2 and the set of $i>2$, etc.: finally $\forall i$, either $x_j>x_i$ $\forall j>i$ or $x_j<x_i\ \forall j>i$. So one of the alternatives holds infinitely often. Take the corresponding subsequence. *Q.E.D.*

PROOF: For $v \neq \lambda u$ we have $S_{u^\varepsilon,v}=\pm u^\varepsilon$ if $\xi_{u^\varepsilon}>\xi_v$, and $S_{u^\varepsilon,v}=\pm v$ if $\xi_{u^\varepsilon}<\xi_v$. For a subsequence as in the claim, one of both cases holds for all small ε, or $\xi_{u^\varepsilon}=\xi_v\ \forall\varepsilon$. Extracting further subsequences to make signs constant, we conclude in the first case from CONT that $S_{u,v}=\lambda u$, so either $\xi_v<\xi_u$ or $\xi_v=\underline{x}$, and in the second that $S_{u,v}=\lambda v$, so either $\xi_u<\xi_v$ or $\xi_u=\underline{x}$—and then too $\xi_u<\xi_v$, since $\xi_v>\xi_{u^\varepsilon}$.

Now, if ξ_{u^ε} is strictly increasing it can be taken $>\underline{x}$; with $v=u^\varepsilon$ we get then $\xi_u>\xi_{u^\varepsilon}$, $\forall\varepsilon$. And if $\xi_u \geq \xi_v \geq \xi_{u^\varepsilon}\ \forall\varepsilon$, similarly $v=\pm u$, so by III(γ) $\xi_v=\xi_u$, thus α.

If ξ_{u^ε} is strictly decreasing, $v=u^\varepsilon$ yields $\xi_u<\xi_{u^\varepsilon}$, $\forall\varepsilon$. And if $\xi_u\le\xi_v<\xi_{u^\varepsilon}$ $\forall\varepsilon$, either $\xi_v=\underline{x}$ or $v=\pm u$. So β is clear if $\xi_u=\inf\xi_{u^\varepsilon}$, and otherwise $\inf_\varepsilon\xi_{u^\varepsilon}=\xi_{-u}>\xi_u$; hence $\xi_u=\underline{x}$ by III(γ), and thus γ.

Take finally $\xi_{u^\varepsilon}=x_0$ $\forall\varepsilon$. For $\xi_u\ne x_0$ (otherwise δ), if $\xi_u=\underline{x}$, then $v\ne -u$, $\xi_v<x_0\Rightarrow\xi_v=\underline{x}$ (i.e., ζ), and otherwise, since $\xi_u\ge\xi_v>x_0$ or $\xi_u\le\xi_v<x_0$ imply $v=\pm u$, III(γ) yields $]x_0,\xi_u[=]\xi_u,x_0[=\varnothing$ and $\xi_v=\xi_u\Rightarrow v=\pm u$. Further $x_0\ne\underline{x}$ (so ϵ), otherwise for $v\ne\pm u$ and $\xi_v=\underline{x}$ we have $S_{u^\varepsilon,v}=0$ $\forall\varepsilon$ (III(α)); hence by CONT $S_{u,v}=0$, so $\xi_u=\underline{x}$ by III(α) and III(δ). The convergence properties of F only remain to prove.

If false, take a subsequence where F_{u^ε} stays away from its stated limit. Also $x_0\ne\underline{x}$, lest the result be obvious (case δ). Make (Claim) F_{u^ε} monotone and with constant sign. Let $v=u^{\varepsilon_0}$ (so $v\ne\pm u$), and make $\varsigma_{u^\varepsilon,v}$ constant: $S_{u^\varepsilon,v}$ is up to sign a convex combination of u^ε and v, whose weights converge (monotonicity of F_{u^ε}). So any monotone subsequence converges specially, to $\varsigma_{u^\varepsilon,v}\operatorname{sign}(F_{u^\varepsilon})u$ if $F_{u^\varepsilon}\to\pm\infty$, and otherwise to $\varsigma_{u^\varepsilon,v}[(\lim F_{u^\varepsilon})u+F_vv]$. Thus, by CONT, $S_{u,v}\in\{0,\varsigma_{u^\varepsilon,v}\operatorname{sign}(F_{u^\varepsilon})u\}$ if $F_{u^\varepsilon}\to\pm\infty$, and otherwise $S_{u,v}\in\{0,\varsigma_{u^\varepsilon,v}[(\lim F_{u^\varepsilon})u+F_vv]\}$. And (III($\alpha$), III($\delta$)) $S_{u,v}\ne 0$ since $\xi_v=x_0\ne\underline{x}$. Thus $\xi_u>\xi_v=x_0$ and $F_vF_{u^\varepsilon}\varsigma_{u,v}\varsigma_{u^\varepsilon,v}>0$ if $F_{u^\varepsilon}\to\pm\infty$, and $\xi_u<x_0$ and $\varsigma_{u,v}=\varsigma_{u^\varepsilon,v}$ if $F_{u^\varepsilon}\to 0$. Otherwise $\xi_u=x_0$, and $\varsigma_{u^\varepsilon,v}F_v=\varsigma_{u,v}F_v$, so $\varsigma_{u,v}=\varsigma_{u^\varepsilon,v}$, and hence $\lim F_{u^\varepsilon}=F_u$, a contradiction. *Q.E.D.*

PART V: *Lemma 1 follows from Steps E and G below.*

Step A

If $\xi_{u^{\pi_1}}=\xi_{v^{\pi_1}}\ne\underline{x}$, $v\ne\pm u$, then $\forall\pi\,F_{v^\pi}/F_{v^{\pi_1}}=F_{u^\pi}/F_{u^{\pi_1}}$, $\xi_{u^\pi}=\xi_{v^\pi}\ne\underline{x}$.

PROOF: $\varsigma_{u^\pi,v^\pi}[uF_{u^\pi}\mathbb{1}_{\xi_{u^\pi}\ge\xi_{v^\pi}}+vF_{v^\pi}\mathbb{1}_{\xi_{v^\pi}\ge\xi_{u^\pi}}]$ is (III(α) and IRA) constant in π on A'. Since $d_{u,v}=2$, both coefficients being nonzero for π_1 implies that for all π, and the constancy of their ratio. *Q.E.D.*

Step B

Fix u such that $\exists w\ne\pm u$: $\xi_w=\xi_u\ne\underline{x}$, and such that $0<u_{a_0}<1$. $\exists\varepsilon>0$ such that if $v_a=u_a$ for $a\ne a_0$ and $|v_{a_0}-u_{a_0}|\le\varepsilon$, then $\xi_v=\xi_u$ and $F_v=F_u$.

PROOF: Fix a_1,a_2: $u_{a_1}<u_{a_0}<u_{a_2}$ and $a_3\ne a_0,a_1,a_2$: $u_{a_1}\le u_{a_3}\le u_{a_2}$ ($\#A\ge 4$). By the Part IV Corollary $\exists v\ne u$: $\xi_v=\xi_u$, $v_a=u_a$ for $a\ne a_3$. For $A'=\complement\{a_0\}$ let $\pi_1\in\Delta^{\{a_1,a_2\}}$ represent u_{a_0}: $v^{\pi_1}=v$ and $u^{\pi_1}=u$, so $\xi_{v^{\pi_1}}=\xi_{u^{\pi_1}}$. Thus if $v^\pi=v^\sigma$, then by A $\xi_{u^\pi}=\xi_{v^\pi}=\xi_{v^\sigma}=\xi_{u^\sigma}$ and $F_{u^\pi}=F_{u^\sigma}$. For σ interior in $P=\Delta^{\{a_1,a_2,a_3\}}$ such that $u=u^\sigma$, $I=\{\pi\in P\mid v^\pi=v^\sigma\}$ is a segment, with σ interior and u^π nonconstant. So $\{u^\pi_{a_0}\mid\pi\in I\}$ is also one, with $u_{a_0}=u^\sigma_{a_0}$ interior and (F_{u^π},ξ_{u^π}) constant. *Q.E.D.*

Step C

If $\exists w\ne\pm u$: $\xi_w=\xi_u\ne\underline{x}$ and if $v_a=u_a$ for $a\ne a_0$, then $\xi_v=\xi_u$, $F_v=F_u$.

PROOF: If $v\ne u$, $\sup_{a\ne a_0}u_a=1$, $\inf_{a\ne a_0}u_a=0$, so by the Part IV Corollary $\xi_v=\xi_u$ for v_{a_0} in an open (in $[0,1]$) interval I around u_{a_0}. Applying B to all $v_{a_0}\in I\setminus\{0,1\}$, F_v is locally constant on $I\setminus\{0,1\}$, hence constant there; thus by IV(δ) $F_v=F_u$ $\forall v\in I$. So $V=\{v_{a_0}\in[0,1]\mid\xi_v=\xi_u$ and $F_v=F_u\}$ is a neighborhood of itself, i.e. open in $[0,1]$. F and ξ being constant on V, IV(δ) implies V is also closed: $V=[0,1]$. *Q.E.D.*

Step D

If $\xi_u \neq \underline{x}$ $\exists \bar{u} \exists v \neq \pm \bar{u}$: $\xi_{\bar{u}} = \xi_v \neq \underline{x}$ such that u and $\bar{u}$ differ in at most two coordinates, which are for each inside the range of the others.

PROOF: Fix $a_1 \neq a_2$ such that u_{a_1} and u_{a_2} are in the range of the others ($\#A \geq 4$). For $y, z \in [0,1]$ let $0 < \varepsilon < \frac{1}{2}$, $u_a^{y,z} = \varepsilon + (1-2\varepsilon)y$ if $a = a_1$, $= \varepsilon + (1-2\varepsilon)z$ if $a = a_2$, $= u_a$ otherwise. Let $\varphi(y,z) = \xi_{u^{y,z}}$: by Part I, $R_{u^{y,z}} = R_u$ yields $\varphi(y,z) \neq \underline{x}$ $\forall y, z$. The $u^{y,z}$ being never equal or opposite to each other, if the claim were false φ would be injective from $[0,1]^2$ to $X \setminus \{\underline{x}\}$. Then, for y going monotonically to y_0, cases γ to ζ are excluded in IV: φ is separately continuous, with the order topology on X. So $P_z^{y_0} = \{y | \varphi(y,z) > \varphi(y_0, z)\}$ is open, and $P_z^{y_0} \cup \{y_0\} = \{y | \varphi(y,z) \geq \varphi(y_0,z)\}$ closed: $P_z^{y_0} \in \{\varnothing, [0, y_0[,]y_0, 1], [0,1] \setminus \{y_0\}\}$. Thus $\varphi(\cdot, z)$ is monotone: otherwise, e.g., $y_1 < y_2 < y_0$, $\varphi(y_1, z) < \varphi(y_0, z) < \varphi(y_2, z)$, a contradiction. $\{z | \varphi(0,z) < \varphi(1,z)\}$ and its complement being open by continuity, either $\varphi(\cdot, z)$ is increasing $\forall z$ or decreasing, and dually for $\varphi(y, \cdot)$: reversing if needed the order on y or z, φ is increasing. Now separate continuity yields continuity, so the sets above and below $(\frac{1}{2}, \frac{1}{2})$ contradict the connectedness of $[0,1]^2 \setminus \{(\frac{1}{2}, \frac{1}{2})\}$.[39] *Q.E.D.*

Step E

$S_u = u$ for all u. In particular, $\underline{x}$ does not exist.

PROOF: Pick u with (I(γ)) $\xi_u \neq \underline{x}$, such that, by D, $\exists v \neq \pm u$: $\xi_v = \xi_u$. Use C at most twice to get $R_u = [0,1]$, keeping $\xi_u \neq \underline{x}$: $\lambda([0,1]) \neq 0$ implies by I(β) $S_u = u$ $\forall u$, lest $S_u = -u$ $\forall u$, which contradicts NOILL. *Q.E.D.*

Step F

$\xi_u = \xi_v$ and $F_u = F_v$ if u and v differ only in finitely many coordinates.

PROOF: Given u, pick $\bar{u}$ by D. By C, $\xi_u = \xi_{\bar{u}} = \xi_v$, so $\forall u \exists v \neq \pm u$: $\xi_v = \xi_u$. Thus $\forall u, v, a_0$ $u_a = v_a$ $\forall a \neq a_0 \Rightarrow \xi_u = \xi_v \& F_u = F_v$ by C. Fix $a_1 \neq a_2$ such that $\sup_{a \neq a_1} u_a = \sup_{a \neq a_1} v_a = 1$. Bring first u_{a_1} to 0, then u_{a_2} to 1, next the finitely many different values of u to v outside $\{a_1, a_2\}$, finally the value at a_2, and then at a_1 to those of v. Each step preserves 0–1-normalization, so by the above ξ and F. *Q.E.D.*

Step G

$S_{u,v} = u + v$ whenever $d_{u,v} = 2$.

PROOF: Fix $A' = \{a_1, a_2, a_3\}$ and $w = \mathbb{1}_{a_1}$. To show that any $\bar{u}$ has the same (ξ, F) as w, assume by F that $u_{a_1} = u_{a_2} = 1$, $u_{a_3} = 0$ for $u = \bar{u}|_{A'}$. For π_1 on $\{a_1, a_3\}$, u^{π_1} and w^{π_1} have same (ξ, F) by F, hence (A) $\xi_{u^\pi} = \xi_{w^\pi}$ and $F_{u^\pi} = F_{w^\pi} \forall \pi$: since for π on $\{a_2, a_3\}$ such that $\bar{u} = u^\pi$, also $w = w^\pi$, ξ and F are constant. Then $s_{u,v} = \operatorname{sign}(F)$ $\forall u, v$ by E and II.A(γ), hence the result by III(α). *Q.E.D.*

LEMMA 2: *Assume $d_{\vec{u}} > 1$ and that RU holds for all $v_{\vec{u}}$-profiles with dimension $d_{\vec{u}} + 1$. If all u_n are in the span of U and S, the profile is exceptional.*

[39] This includes the proof of the lemma: No total order on (a region with nonempty interior of) $\mathbb{R}^n (n > 1)$ is separately continuous (i.e., such that $\forall p_1, p_2, x_{-i}$, $\{x_i | p_1 < (x_i, x_{-i}) < p_2\}$ is open).

PROOF: Clearly $d_{\vec{u}} = 2$, so $U \neq 0$, $S \notin \{0, U\}$, hence $\nu_{\vec{u}} \geq 3$ by Lemma 1. Let $u_n = \alpha_n U + \beta_n S + \gamma_n$.

Claim

$\forall n_0 \in N_+$ and $\forall a_0$ one of the following holds:
(α) $U_a = \lambda_u v_a + \mu_u$, $S_a = \lambda_s v_a + \mu_s$ $\forall a \neq a_0$ (for some $\lambda_u, \mu_u, \lambda_s, \mu_s$, and v);
(β) $\alpha_n = \alpha t_n$, $\beta_n = \beta t_n$ $\forall n \neq n_0$ (for some α, β, t);
(γ) $U_{|C\{a_0\}}$ is nonconstant and $U(a_0) \notin [\inf_{a \neq a_0} U(a), \sup_{a \neq a_0} U(a)]$;
(δ) $u_{n_0|C\{a_0\}}$ is nonconstant and $u_{n_0}(a_0) \notin [\inf_{a \neq a_0} u_{n_0}(a), \sup_{a \neq a_0} u_{n_0}(a)]$;
(ϵ) $U(a_0) = \sup_{a \neq a_0} U(a)$ and $u_{n_0}(a_0) = \inf_{a \neq a_0} u_{n_0}(a)$;
(ζ) $U(a_0) = \inf_{a \neq a_0} U(a)$ and $u_{n_0}(a_0) = \sup_{a \neq a_0} u_{n_0}(a)$.

PROOF: Assume α to ζ false for (n_0, a_0) with $n_0 \in N_+$. γ to ζ being so, if $u^{\varepsilon}_{n_0}(a_0) = u_{n_0}(a_0) + \varepsilon$, $u^{\varepsilon}_{n,a} = u_{n,a}$ for $(n, a) \neq (n_0, a_0)$, and $U^{\varepsilon}(a) = \sum_n u^{\varepsilon}_n(a)$, then $u^{\varepsilon}_{n_0}$ and U^{ε} converge specially, either for $\varepsilon > 0$ or for $\varepsilon < 0$ (recall U^{ε} always does if U is constant for $a \neq a_0$). The special limit of $u^{\varepsilon}_{n_0}$ is u_{n_0}, since $n_0 \in N_+$, and that of U^{ε} is U, since $U \neq 0$. Finally $d_{\vec{u}^{\varepsilon}} = 3$, i.e., for some $\underline{a} \neq a_0$ we need

$$\operatorname{rank}\begin{pmatrix} u_{n_0}(a_0) + \varepsilon - u_{n_0}(\underline{a}) & \cdots & u_n(a_0) - u_n(\underline{a}) \\ \vdots & & \vdots \\ u_{n_0}(a) - u_{n_0}(\underline{a}) & \cdots & u_n(a) - u_n(\underline{a}) \end{pmatrix} = 3.$$

So some 3×3 subdeterminant must be nonzero. It has to contain the entry (n_0, a_0), since $d_{\vec{u}} = 2$. So it is $(n_0, n_1, n_2) \times (a_0, a_1, a_2)$, and for the same reason it equals zero for $\varepsilon = 0$, so the coefficient of ε

$$\begin{vmatrix} u_{n_1}(a_1) - u_{n_1}(\underline{a}) & u_{n_2}(a_1) - u_{n_2}(\underline{a}) \\ u_{n_1}(a_2) - u_{n_1}(\underline{a}) & u_{n_2}(a_2) - u_{n_2}(\underline{a}) \end{vmatrix}$$

must be nonzero. Otherwise $(u_n(a) - u_n(\underline{a}))_{n \neq n_0, a \neq a_0}$ has rank ≤ 1, i.e., $\forall n \neq n_0$, $\forall a \neq a_0$, $u_n(a) - u_n(\underline{a}) = \lambda_n w(a)$, so $u_n(a) = \lambda_n w(a) + \mu_n$. Since $u_n(a) = \alpha_n U_a + \beta_n S_a + \gamma_n, n$, we get

$$\alpha_n U_a + \beta_n S_a = \lambda_n w_a + \nu_n \qquad \forall_n \neq n_0, \forall a \neq a_0.$$

α being false, $\dim\{(U_a, S_a) | a \neq a_0\} = 2$, so $\exists a_1, a_2, \underline{a} \in C\{a_0\}$ such that

$$\begin{vmatrix} U_{a_1} - U_{\underline{a}} & S_{a_1} - S_{\underline{a}} \\ U_{a_2} - U_{\underline{a}} & S_{a_2} - S_{\underline{a}} \end{vmatrix} \neq 0.$$

Then $\forall n \neq n_0$ the solution of $\alpha_n[U_{a_i} - U_{\underline{a}}] + \beta_n[S_{a_i} - S_{\underline{a}}] = \lambda_n[w_{a_i} - w_{\underline{a}}]$ $(i = 1, 2)$ is unique and of the form $\alpha_n = \lambda_n F(a_1, a_2, \underline{a})$, $\beta_n = \lambda_n G(a_1, a_2, \underline{a})$; thus β, a contradiction.

$d_{\vec{u}^{\varepsilon}} = 3$ implies $\preccurlyeq^{\varepsilon}_{n_0} \neq \preccurlyeq_{u_0}$, so $(u_{n_0}(a))_{a \neq a_0}$ is nonconstant and hence, δ being false, $u^{\varepsilon}_{n_0}$ is still 0–1-normalized. U^{ε} represents thus the social preferences for $\vec{u}^{\varepsilon}$. So $S \in \{0, U\}$ by CONT, a contradiction. *Q.E.D.*

It remains to show that a profile satisfying the claim is exceptional.

(a) Let $N_{\beta} = \{n_0 \in N | (\beta) \text{ holds}\}$. Then $\# N_{\beta} \leq 1$. Otherwise $\alpha_n = \alpha^i t^i_n$, $\beta_n + \beta^i t^i_n$ $\forall n \neq n_i$, for $n_i \in N_{\beta}$, $i = 1, 2$. By $\nu_{\vec{u}} \geq 3$, $\exists \bar{n} \notin \{n_1, n_2\}$: $(\alpha_{\bar{n}}, \beta_{\bar{n}}) \neq (0, 0)$. Say $\alpha_{\bar{n}} \neq 0$, so $\alpha^i \neq 0$, $t^i_n = \alpha_n / \alpha^i$ $\forall n \neq n^i$, and $\beta_n = \alpha_n(\beta^i / \alpha^i)$ $\forall n \neq n_i$, thus $\beta^i / \alpha^i = \beta_{\bar{n}} / \alpha_{\bar{n}}$, $= \lambda$ say: $\beta_n = \lambda \alpha_n$ $\forall n$, so $d_{\vec{u}} \leq 1$, a contradiction.

(b) $A = A_{\gamma} \cup A_{\delta}$ for $A_{\gamma} = \{a_0 \in A | \gamma \text{ holds}\}$, $A_{\delta} = \{a_0 \in A | \exists n \notin N_{\beta} \cup N_0\text{: } \delta \text{ holds at } (a_0, n)\}$: Any $a_0 \notin A_{\gamma} \cup A_{\delta}$ is in $A_{\alpha} = \{a_0 \in A | \alpha \text{ holds}\}$. Otherwise ϵ or ζ hold $\forall n_0 \notin N_{\beta} \cup N_0$, and the same, say ϵ, lest U be constant. So $u_n(a_0) = 0$ $\forall n \notin N_{\beta}$. Thus $N_{\beta} = \{\underline{n}\}$ by (a); otherwise $U_{a_0} = 0$ yields $U_a = 0$ $\forall a(\epsilon)$. Then the $u_n = t_n[\alpha(U - U_{a_0}) + \beta(S - S_{a_0})]$ $\forall n \neq \underline{n}$ are equal (normalization), so $2 \leq \nu_{\vec{u}} - 1 \leq \sup U_a \leq U_{a_0} \leq 1$.

If $a_1 \in A_\alpha$, all u_n are of the form $\delta_n v_a + \varepsilon_n$ $\forall a \neq a_1$, so U too. Thus by the above v_a attains at any $a \notin A_\alpha$ either its single maximum or its single minimum over $C\{a_1\}$: there being at most two such points, $\exists a_2 \neq a_1$ in A_α, with $u_n(a) = \delta'_n v'_a + \varepsilon'_n$ $\forall n, a \neq a_2$. Then $u_n(a)$ is constant on $C\{a_1, a_2\}$ $\forall n$. Otherwise, if e.g., u_1 is not, express v and v' in terms of u_1: $\forall n$ $u_n(a) = \delta_n u_1(a) + \varepsilon_n$ $\forall a \neq a_1$, $= \delta'_n u_1(a) + \varepsilon'_n \forall a \neq a_2$; u_1 being nonconstant on $C\{a_1, a_2\}$ implies $\delta'_n = \delta_n$ and $\varepsilon'_n = \varepsilon_n$, so $u_n(a) = \delta_n u_1(a) + \varepsilon_n$ $\forall a, n$, contradicting $d_{\vec{u}} > 1$. No u_n having thus a single maximum or minimum on $C\{a_1, a_2\}$, $C\{a_1, a_2\} \subseteq A_\alpha$ (cf. supra) $A_\alpha = A$. Then any u_n is constant outside any pair, hence constant: a contradiction again, so $A_\alpha = \varnothing$.

(c) If $A_\gamma = \varnothing$, then $u_n(a) \in \{0,1\}$ $\forall a, n \notin N_\beta$. So δ never applies (or else u_{n_0} would be constant for $a \neq a_0$), contradicting (b). If $A_\gamma = \{a_0\}$, e.g. with $U_{a_0} < \inf_{a \neq a_0} U_a$, ζ cannot occur. Hence $\forall n \notin N_0 \cup N_\beta$, $\forall a \neq a_0$, $u_n(a) \in \{0,1\}$, with a single "1" for each n. So for $a \neq a_0$, if $\exists n \notin N_0 \cup N_\beta$: $u_n(a) = 0$, then (ϵ) $U_a = \sup U$, and therefore (a fortiori) also otherwise, contradicting $a_0 \in A_\gamma$. So $\# A_\gamma = 2$ ($\# A_\gamma \leq 2$ is clear). Hence ϵ and ζ never hold; thus $\# A_\delta \leq 2$, so $= 2$ by (b) since $\# A \geq 4$. The conditions of Definition 2 follow immediately. *Q.E.D.*

LEMMA 3: *If RU holds for $(d, \nu-1)$-, $(d-1, \nu-1)$-, and nonexceptional $(d+1, \nu)$-profiles, then:*
(α) *for nonexceptional (d, ν)-profiles with $d \neq 1$ we have $S = U$,*
(β) *for exceptional (d, ν)-profiles $S = \sum_n \lambda_n u_n$, $S \neq 0$, $S \neq -U$, $U \neq 0$;*
(γ) *if $d = 1$, then $S \in \{0, U\}$.*

PROOF: Assume $\nu \geq 2$ by Lemma 1, so $d \geq 1$. If $U \neq 0$ and $d > 1$, RU for $\nu - 1$ and Proposition 1.1 yield $\forall n \in N_+$ $\alpha_n S = \mu_n u_n + \tau_n(U - u_n)$. If $\exists n$: $\mu_n = \tau_n$, take by Proposition 1.1 $\mu_n = \tau_n = 1$ and $\alpha_n \leq 0$, so $S = U$ since $U \neq 0$. Otherwise, all u_n being in the span of S and U (trivial if $n \in N_0$), $\vec{u}$ is exceptional by Lemma 2 ($d > 1$ prevents exceptional $(d+1)$-profiles), so $d_{\vec{u}} = 2$ implies then S is also in the span of the u_n, and $d_{U,S} = 2$. Thus α in case $U \neq 0$; β follows also, since for exceptional profiles $U \neq 0$ and $d = 2$.

It remains to prove α for $U = 0$, and γ. If $U \neq 0$, take n_0 in the majority. For $\inf_{a \neq a_0} u_{n_0}(a) \leq u_{n_0}(a_0) \leq \sup_{a \neq a_0} u_{n_0}(a)$, choose $u^\varepsilon_{n_0}(a_0) = u_{n_0}(a_0) + \varepsilon \in [0,1]$, $\vec{u}^\varepsilon = \vec{u}$ elsewhere. Then $d_{\vec{u}^\varepsilon} > 1$: if $d_{\vec{u}} = 1$, since $\nu \geq 2$; the other case is obvious. Further $U^\varepsilon \neq 0$, and $\vec{u}^\varepsilon$ is not exceptional: if $U = 0$, because U^ε is two-valued; otherwise $d_{\vec{u}} = 1$, hence $u_n = \pm U$ for $n \notin N_0 \cup \{n_0\}$, so if $U = u_n$ has a single maximum and a single minimum, they are also of the sum of utilities U^ε, for ε small. Thus by the previous case $\vec{u}^\varepsilon$ satisfies RU.

Use CONT: if $U = 0$, U^ε converges specially to $\operatorname{sign}(\varepsilon)\mathbb{1}_{a_0}$, so if $S \neq U$ $(=0)$ then $S = \operatorname{sign}(\varepsilon)\mathbb{1}_{a_0}$; this yields a contradiction by changing a_0 ($\# A \geq 4$). Otherwise U^ε converges specially to U, so $S \in \{0, U\}$. *Q.E.D.*

PROOF OF THE PROPOSITION: If $d_{\vec{u}} > 2$, clearly $\# N \geq 3$ and $\# A \geq 4$, so it suffices to prove α to δ, and hence, with $i = \nu - d$, that if they hold when $i < i_{\vec{u}}$ or $i = i_{\vec{u}}$, $d < d_{\vec{u}}$, they hold for $\vec{u}$. Then by induction, if they hold for i_0 with $d = 0$, they do so $\forall d$, thus for $i_0 + 1$ and $d = 0$. Being vacuously true for $i = d = 0$, they hold then $\forall i, d$ by induction on i. So by Lemma 3, what remains to prove is a contradiction when $\nu_{\vec{u}} < \# N$ and:

(a) $d_{\vec{u}} = 1$, $S = 0 \neq U$; or

(b) $\vec{u}$ is exceptional and $d_{U,S} = 2$.

Choose $v \notin L$, the span of the u_n, such that, if all but one $u_n (n \in N_+)$ equal w or $-w$, with no less of the former as of the latter, then $\sup_a(v_a + w_a) = 2$, $\inf_a(v_a + w_a) = 0$. Set u_{n_0} to v for some $n_0 \in N_0$.

$\forall n \in N_+$, the profile $(\vec{u}_{-n}, v)$ satisfies RU: if (a), by Lemma 1 if $\nu_{\vec{u}} < 2$, otherwise the dimension being 2; if (b), when n is the exceptional individual, the sum of utilities is maximal or minimal with w, so the profile is not exceptional and RU was obtained for it above, in the call to Lemma 3. When not, $d_{\vec{u}_{-n}} = 2$, so $v \notin L$ yields $d_{\vec{u}_{-n}, v} = 3$. Thus $S^*_n = S_{\vec{u}_{-n}, v} = U + v - u_n$, so $v \notin L$ and $u_n \neq 0$ imply $d_{u_n, S^*_n} = 2$: by Proposition 1.3, $S_{\vec{u}, v} = \mu_n u_n + \tau_n(U + v - u_n) \neq 0$, and $= \mu v + \tau S$ since $v \notin L$ and $S \in L$.

In case (a), $S = 0 \neq S_{\vec{u},v}$ yields thus $S_{\vec{u},v} = \pm v$. So the profile is spanned by $U + v$ and $S_{\vec{u},v}$, hence it is exceptional by Lemma 2, which is impossible by the choice of v.

In case (b), $\tau_n = \mu$ since $v \notin L$ while $S, U \in L$. Hence $\rho_n u_n = \tau S - \mu U$ with $\rho_n = \mu_n - \mu$, so $d_{\vec{u}} = d_{U,S} = 2 \Rightarrow \mu = \tau = 0 \Rightarrow S_{\vec{u},v} = 0$: a contradiction. *Q.E.D.*

APPENDIX C: NECESSITY OF MON1–MON4

The following examples show that each of MON1–MON4 is necessary:

MON1: Let $S = \sum_n \lambda_n^* u_n$, with $\lambda_n^* = 1 - \varepsilon x_n^*$, where $x^* \in \overline{R}_{\vec{u}}$ is such that $\langle \lambda^*, x^* \rangle \geq \langle \lambda^*, x \rangle$ $\forall x \in R_{\vec{u}}$, i.e., λ^* is the gradient of $-\sum_n (\varepsilon^{-1} - x_n)^2$ at its maximizing point x^* over $\overline{R}_{\vec{u}}$. Viewing utility functions as points in the quotient of $L_2(\nu)$ by the constant functions, with ν normalized Lebesgue measure on $R_{\vec{u}}$, $\varepsilon = \min((\gamma/\#N)\sqrt{1-\mu}, \sqrt{\delta})$ where $\gamma = \|U\|$,

$$\mu = \max_n \max\left(\frac{\langle U, u_n \rangle}{\|U\| \, \|u_n\|}, \frac{\langle U, U - u_n \rangle}{\|U\| \, \|U - u_n\|} \right)$$

and δ is the kth eigenvalue, in decreasing order, of the covariance matrix of the $(u_n)_{n \in N}$, with $k = \min(\#A - 1, \#N)$.[40]

MON2: $S = 0$ if $d_{\vec{u}} \leq 1$ and $v_{\vec{u}} \leq k$, $= V$ ("imposed") (or U, or $-U$ if $k \neq 0$) otherwise. Or, to satisfy strict Pareto: $S = U$ if $d_{\vec{u}} > 1$ or unanimity holds, $= 0$ otherwise.[41]

[40] Since $\varepsilon < 1$, strict Pareto holds, and since λ^* depends only on $\overline{R}_{\vec{u}}$, IRA also, in its strongest form, and ANON clearly. As to MON, note that $S = U$ when $v_{\vec{u}} < \#N$; the value of ε ensures that for $d_{u_n, U} = 2$ the angle of S with U is less than those of u_n and of $U - u_n$, so the projection of S on the plane spanned by u_n and $U - u_n$ is a *positive* linear combination of them, hence MON4. And if $\exists n$: $d_{u_n, U} < 2$, then $S = U$, so the strongest form of MON holds. Next, x^* is jointly continuous in ε and $\overline{R}_{\vec{u}}$ (Hausdorff topology) for $0 < \varepsilon < 1$. Thus λ^* is so for $0 \leq \varepsilon < 1$. ε is continuous in the covariance matrix, which depends itself continuously on $\overline{R}_{\vec{u}}$ when $d_{\vec{u}}$ stays constant: since the term $\sqrt{\delta}$ forces $\varepsilon = 0$ in all but the highest dimension, ε and hence λ^* depend continuously on $\overline{R}_{\vec{u}}$. This yields any desired continuity property; since also $u_n = 0 \Rightarrow S = U$ and $S = 0 \Leftrightarrow U = 0$, $\{\vec{u} | S \neq U\}$ is open, using on preferences the quotient of the uniform topology on utility functions. Finally, $S \neq U$ when $\varepsilon > 0$ except if, with x^* the maximizer of U on $\overline{R}_{\vec{u}}$ closest to the diagonal, the affine subspace orthogonal at x^* to $R_{\vec{u}}$ intersects the diagonal. So the sequential closure of $\{S \neq U\}$ under special convergence is $\{\vec{u} | v_{\vec{u}} = \#N\}$.

[41] MON2 is not needed for the aggregation itself, but just to have $S_u = u$: replacing INDIV, NONT, NOILL, and MON2 by this statement, the proposition still holds, adding in α that $d_{\vec{u}} \neq 1$. Indeed: (a) in the proof of II.A, $S^* \neq 0$, (b) in that of Lemma 3, $\alpha_n S = \mu_n u_n + \tau_n (U - u_n)$ still holds when $u_n \neq U$; and otherwise u_n is anyway in the span of S and U; (c) in the proof of the proposition, one no longer needs case (a), but as seen there, if $d_{\vec{u}} = 1$ and $d_{\vec{u},v} = 2$, then $0 \neq S_{\vec{u},v} \in L$, and further $S_{\vec{u},v} \neq -U$ since $\max(\mu, \tau) \geq 0$, so if $S_{\vec{u},v} \neq U$ the profile is spanned by U and $S_{\vec{u},v}$, thus is exceptional by Lemma 2.

In the presentation of MON then, (a) take for MON, weak Pareto and consistency (only the strict inequality part), (b) in Proposition 1 drop MON2, require in 1 that $(\alpha\mu, \tau) \neq (0,0)$ and $\min(\alpha\mu, \alpha\tau) \leq 0$, and adjust 2 and 3, and (c) in footnote 27 change "Just the first part" to "$p \succ^* q \,\&\, p \succ_n q \Rightarrow p \not\prec q$" and in footnote 28 add "or the weak Pareto condition."

This may seem a simpler, more elegant, and possible weaker axiomatization: after all, $S_u = u$ is a very obvious case of Pareto. Its lumping together of axioms with very different (and strong) ethical content, like the negotiation of objective ethical norms (INDIV) and a strictly positive association of social preferences with every single individual's, is no good objection since, e.g., if $\#A < \infty$ it can be further analyzed (cf. Part I) in INDIV, the particular case of MON2 where $u_i = 0$ $\forall i \neq n$, NOILL and the strengthening of NONT to "S_u is not identically zero." Besides the difficulty to do this for infinite A, the main advantage of the present approach is that this strengthening of NONT *is* substantial, requiring social preferences to be affected when a single individual out of millions changes his preferences (cf. fn. 38), while MON2 is fully in the spirit of the rest of MON, and, its above-mentioned particular case being needed anyway, conceptually a very minor strengthening.

MON3: Fix $L > 1$. $S = 0$ if $v_{\vec{u}} = \#N$ and $\exists\lambda$: $\sum_{n \in K} \lambda_n u_n = 0$ with $1 \leq \lambda_n \leq L = U$, otherwise.
MON4: $S = U$ if $v_{\vec{u}} < k$, $= -U$ otherwise.[42]

REFERENCES

Arrow, K. J. (1963): *Social Choice and Individual Values*, Vol. 12 of *Cowles Foundation Monographs*, 2nd edn. New York: John Wiley & Sons, Inc.

d'Aspremont, C. (1985): "Axioms for Social Welfare Orderings," in Hurwicz et al. (1985), pp. 19–76.

Aumann, R. J., and S. Hart (Eds.) (1992): *Handbook of Game Theory* (*with Economic Applications*), Vol. 1, No. 11, in Handbooks in Economics. Amsterdam: North-Holland.

Brams, S. J., and P. C. Fishburn (1978): "Approval Voting," *American Political Science Review*, 72, 831–847.

Chichilnisky, G., and G. Heal (1983): "Necessary and Sufficient Conditions for the Resolution of the Social Choice Paradox," *Journal of Economic Theory*, 31, 68–87.

[42] As shown in the previous footnote, with strict Pareto one can dispense with INDIV, NONT, NOILL, and MON2. Dropping also MON4, the proof still yields, replacing every U by $\pm U$, a similar proposition: α yields $S = \pm U$, only if $d_{\vec{u}} \neq 1$, β stays, γ becomes $U = 0 \Rightarrow S = 0$, and in δ one gets only $S \neq 0$. Now strict Pareto excludes $-U$ and thus forces U in some region. By "continuity" (and convexity of $\{U \neq 0\}$) the result should thus still follow.

This doesn't carry through with the present form of CONT: e.g. if $N = \{0, \ldots, 6\}$, $A = \{a_0, \ldots, a_3\}$, $u_n = \mathbb{1}_{a[\frac{n}{2}]} + x_n \mathbb{1}_{a_0}$, $x_0 = 0$, $x_i \geq 0$, $\sum x_i < 1$, then every u_n has a single 1, so those can't be lowered, and for $a \neq a_0$ the zeros too can't be changed, since the $0-1$-normalized U would change at two different coordinates, hence not be specially continuous. Thus the only possible changes are those of the x_n, and they have to respect the given constraints since for $\sum x_i = 1$ one would have U constant, hence also not specially continuous. So CONT relates such profiles only to other such profiles; since every lottery is Pareto-efficient for any such profile, strict Pareto doesn't have any bite either.

However with the strengthened version of CONT in Section 6, the argument does go through:

First, by Section 6, the proposition yields now that $S \in \{-U, 0, U\}$ and that $d_{\vec{u}} \neq 1 \Rightarrow S = \pm U$. It follows thus in particular from Pareto that $v_{\vec{u}} \leq 2 \Rightarrow S = U$. Take then an arbitrary profile $\vec{u}$ and use the construction in Step F for each individual in turn to join $\vec{u}$ to some other profile, where two fixed alternatives have utilities respectively 0 and 1 for all nonindifferent individuals, by a path of 0–1-normalized profiles $\vec{u}_t$ for $0 \leq t \leq k$ where on each subinterval $[i, i+1]$ a single coordinate of the profile changes (linearly). Then there are only finitely many t's with $d_{\vec{u}_t} = 1$ or $U = 0$ (or both). Make those integers too, reparameterizing the path; doubling t makes them even, so everywhere else on both adjacent segments one has both $d_{\vec{u}_t} > 1$ and $U_t \neq 0$. If at some such critical point one has a_0 (the alternative along the incoming segment) equal to a_1 (along the outgoing segment), insert at that point, between the two segments, a perturbation of $u_n(\bar{a})$, for some n and some $\bar{a} \neq a_0$ such that $\sup_{a \neq \bar{a}} u_n^a = 1$, $\inf_{a \neq \bar{a}} u_n^a = 0$, followed by its reversal. For a sufficiently small perturbation one will have $d_{\vec{u}_t} > 1$ and $U_t \neq 0$ at all additional points, so the new path has all properties of the old one. So we have further always $a_0 \neq a_1$.

If interchanging in a close vicinity of the critical point the change before it with the one after it yields a 0–1-normalized profile at the new intermediate point, then the new path has this critical point less (recall we have $v_{\vec{u}_t} \geq 3$). Otherwise $n_0 = n_1$, $a_0 \neq a_1$, and, with $w = u_{n_0}^t$ at the critical t, one has $w^{a_0} = w^{a_1} = 0 < \inf_{a \notin \{a_0, a_1\}} w^a$, or its dual. Then w can be slightly perturbed at some $a_2 \notin \{a_0, a_1\}$ while remaining 0–1-normalized. Introducing such a perturbation, followed by its reversal, at the critical point yields a new path where the original critical point is duplicated, but all new points are noncritical. And now each duplicate of the critical point satisfies still $a_0 \neq a_1$, but also $w^{a_0} \neq w^{a_1}$, so will disappear by the previous interchange.

Thus $d_{\vec{u}_t} > 1$ and $U_t \neq 0$ at all interior t. By Pareto $S = U \neq 0$ at the endpoint. So by continuity we get first that $S = U$ at all interior t, next that $S \in \{0, U\}$ at the arbitrary initial profile. Thus also $d_{\vec{u}} \neq 1 \Rightarrow S = U$.

Thus MON1 and MON3 form indeed the irreducible core of MON.

DHILLON, A. (1998): "Extended Pareto Rules and Relative Utilitarianism," *Social Choice and Welfare*, forthcoming.

DHILLON, A., AND J.-F. MERTENS (1993): "Relative Utilitarianism," Discussion Paper 9348, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

——— (1997): "An Impossibility Theorem with von Neumann-Morgenstern Preferences," *Economic Letters*, 56, 305–309.

DIAMOND, P. (1967): "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility: A Comment," *Journal of Political Economy*, 75, 765–766.

DRÈZE, J. H. (1981): "Inferring Risk Tolerance from Deductibles in Insurance Contracts," *The Geneva Papers on Risk and Insurance*, 20, 48–52.

HAMMOND, P. J. (1988): "Consequential Foundations for Expected Utility," *Theory and Decision*, 25, 25–78.

HARSANYI, J. C. (1976): *Essays on Ethics, Social Behavior, and Scientific Explanation*, No. 12 in Theory and Decision Library. Dordrecht, Holland: D. Reidel Publishing Company.

——— (1977): *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

——— (1992): "Game and Decision Theoretic Models in Ethics," in Aumann and Hart (1992), Ch. 19, pp. 669–707.

HURWICZ, L., D. SCHMEIDLER, AND H. SONNENSCHEIN (EDS.) (1985): *Social Goods and Social Organization—Essays in Memory of Elisha Pazner*. Cambridge: Cambridge University Press.

KALAÏ, E., AND M. SMORODINSKY (1975): "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43, 307–328.

MAY, K. O. (1952): "A Set of Independent Necessary and Sufficient Conditions for Majority Decision," *Econometrica*, 20, 680–684.

MOULIN, H. (1988): *Axioms of Cooperative Decision Making*, No. 15 in Econometric Society Monographs. Cambridge: Cambridge University Press.

RAWLS, J. (1971): *A Theory of Justice*. London: Oxford University Press.

SEN, A. K. (1970): *Collective Choice and Social Welfare*, No. 5 in Mathematical Economics Texts. San Francisco: Holden-Day.

——— (1973): *On Economic Inequality*. London: Oxford University Press.

References

Approval Voting
Steven J. Brams; Peter C. Fishburn
The American Political Science Review, Vol. 72, No. 3. (Sep., 1978), pp. 831-847.
Stable URL:
http://links.jstor.org/sici?sici=0003-0554%28197809%2972%3A3%3C831%3AAV%3E2.0.CO%3B2-M

Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment
Peter A. Diamond
The Journal of Political Economy, Vol. 75, No. 5. (Oct., 1967), pp. 765-766.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28196710%2975%3A5%3C765%3ACWIEAI%3E2.0.CO%3B2-H

A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision
Kenneth O. May
Econometrica, Vol. 20, No. 4. (Oct., 1952), pp. 680-684.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28195210%2920%3A4%3C680%3AASOINA%3E2.0.CO%3B2-S