Soc Choice Welfare (1994) 11: 331-354

Social Choice and Welfare

Harsanyi's Aggregation Theorem: multi-profile version and unsettled questions

Philippe Mongin*

CORE, Université Catholique de Louvain, 34, Voie du Roman Pays, B-1348 Louvain-La-Neuve, Belgium (Fax: 32-10 47 43 01)

Received December 21, 1992 / Accepted January 25, 1994

Abstract. Harsanyi's Aggregation Theorem states that if the individuals' as well as the moral observer's utility functions are von Neumann-Morgenstern, and a Pareto condition holds, then the latter function is affine in terms of the former. Sen and others have objected to Harsanyi's use of this result as an argument for utilitarianism. The present article proves an analogue of the Aggregation Theorem within the multi-profile formalism of social welfare functionals. This restatement and two closely related results provide a framework in which the theorem can be compared with well-known characterisations of utilitarianism, and its ethical significance can be better appreciated. While several interpretative questions remain unsettled, it is argued that at least one major objection among those raised by Sen has been answered.

1. General

One of the most fascinating results of early social choice theory is Harsanyi's 1955 Aggregation Theorem: Suppose that the individuals' utilities u_i as well as the social planner's or moral observer's utility u are von Neumann-Morgenstern, and the Pareto-indifference rule holds; then the latter function is affine in terms of the former, i.e., $u = \sum a_i u_i + b$. Harsanyi has repeatedly used this result as an axiomatic justification for utilitarianism. It is one of the three main arguments that he provides in support of this theory, the two others being his well-known Impartial Observer Theorem (1953), which is Harsanyi's version of the "veil of ignorance", and a direct, nonaxiomatic argument for the factual basis of interpersonal comparisons of the utilitarian sort (to be found, e.g., in 1955, Sect. 5).

* The author is grateful to V. Barham, J. Broome, M. Fleurbaey, D. Hausman, S. Kolm, J. Roemer, and P. Suppes, for useful discussions and suggestions. Special thanks are due to C. d'Aspremont, N. McClennen, J. Weymark, and an anonymous referee for detailed comments on an earlier version. The usual caveat applies. The author also gratefully acknowledges financial support from the SPES programme of the Union Européenne.

There are delicate interconnections between these three pieces of doctrine, but each can be considered separately.[1] Although the 1955 Aggregation Theorem should be of much interest to theoretical economists, relying as it does on very standard assumptions of their field, its significance is not as easily understood as that of the other two arguments. As Harsanyi himself has noted, "[the Aggregation Theorem] yields a lesser amount of philosophically interesting information about the nature of morality than [the Impartial Observer Theorem], but it has the advantage of being based on much weaker – almost trivial – philosophical assumptions. [...] It relies merely on Pareto-optimality and on the Bayesian rationality postulates" (1977b, in Sen and Williams, 1982, p 48).

Both the technical aspects and the significance of the Aggregation Theorem have been actively debated. Even leaving aside the former issue[2], there is much disagreement amongst the discussants. Hammond's comments (1982, 1987, 1991) can be read as partially supportive ones. The consequentialist reconstruction of utilitarianism that Hammond promotes appears to share significant common features with Harsanyi's approach in the Aggregation Theorem. The same applies to Myerson's (1981) discussion of the timing effect of collective decision. Sen is prominent among the opponents. The better-known part of the Sen-Harsanyi debate, as it has recently come to be called (Weymark 1991), relates to the distributive consequences of utilitarianism and has therefore little to do with the Aggregation Theorem per se.[3] The other part, on which Weymark (1991) usefully expands, revolves around the following claim made by Sen: "there is a need for an axiomatic derivation of utilitarianism despite Harsanyi's theorems" (1986, p 1124). As far as the Aggregation Theorem is concerned, Sen submits that it is "more assertive" (ibid.) than the Impartial Observer Theorem, but still suffers from various conceptual defects. To the best of our knowledge, this strand of criticism has not provoked any answer from either Harsanyi himself or another sympathetic reader. Most noteworthy among the other discussions of the Aggregation Theorem are Broome's (1987, 1990, 1991). At the start of a careful analysis of the philosophical consequences of the result, he writes: "It would be remarkable if a formal argument could establish a moral theory. This one does not; but I think it can contribute to our understanding of utilitarianism" (1987, p 405).

The conclusion of this paper is in close accord with this highly qualified endorsement of Harsanyi's approach in his Aggregation Theorem. We also believe

[1] In this later work (e.g., 1977a), Harsanyi tends to use the three construals together as mutually supporting arguments. This occasionally results in lack of clarity. As will be seen below, the significance of the Aggregation Theorem would be weakened if it required a separate argument to justify interpersonal comparisons of the utilitarian sort. Note that some readers appear not to distinguish the Aggregation Theorem from the Impartial Observer Theorem, a very unfortunate confusion.

[2] Many proofs of the (single-profile) Aggregation Theorem have been published and some interesting corollaries added to it. Coulhon and Mongin (1989) cover this strand of discussion while providing their own quick proof. See also Weymark (1991, 1994) and the bibliography in these papers. *Subjective uncertainty* variants of the Aggregation Theorem have recently come into the open: Broome's (1990) relies on Jeffrey's axiomatisation of Bayesianism, whereas Mongin's (1993) depends on Savage's axioms. It is impossible to derive a subjective uncertainty variant of the Aggregation Theorem without making use of *some* axiomatic framework for Bayesianism: see Mongin (1993, § 4.1).

[3] This line of criticism started with Diamond (1967) and is elaborated upon in Sen, 1970, p 142–145; see also Harsanyi's 1975 response.

that the latter achieves something significant, although its ethical message is not easily summarised. However, we shall not attempt to offer a complete reconstruction here. Our main business is to prove the *multi-profile* analogue of the Aggregation Theorem that is noticeably missing in contemporary social choice theory. For the main, we shall limit our interpretative contribution to the arguments on which the multi-profile approach casts light.

Interestingly, the lack of a multi-profile equivalent was part of Sen's objections to Harsanyi's approach. Let us briefly review the reasons for which he claimed that the Impartial Observer and the Aggregation Theorems are not proper axiomatisations of utilitarianism. "What is needed is an axiomatisation that (1) permits independent formulation of individual utilities, and (2) which has the invariance property of having the set of a_i [= the weights given to the individuals in the utility sum] determined independently of the utility functions to be aggregated (and in particular having $a_i = 1$)" (1986, p 1124). Reason (1) involves subtle objections which are clarified and endorsed by Weymark (1991). We shall return to it in the concluding section of this paper and meanwhile concentrate on reason (2), which is easy to grasp and will be shown to be amenable to a clear-cut answer. Sen's point here appears to be that the standard Aggregation Theorem takes the $(n+1)$-utilities of the individuals and the social observer as given, and therefore makes the weights a_i dependent on the chosen profile[4]. That is, Harsanyi follows the traditional Bergson-Samuelson approach to the "social welfare function." We agree with Sen that this is inappropriate for an axiomatisation of utilitarianism. One major reason that this is so is just alluded to at the end of Sen's quotation and can be made explicit as follows. In contrast with affine or weighted sum rules, which have been rarely defended in the tradition of ethics and political philosophy, *classical utilitarianism* always involves uniform weights, be they equated to 1 – as in Bentham's sum rule – or to $\frac{1}{n}$ – as in the (also time-honoured) average utility rule, Harsanyi's own favourite. To go from the affine conclusion of the Aggregation Theorem to classical utilitarianism one must satisfy a symmetry requirement which cannot, strictly speaking, be stated in the language of single profiles. This argument needs some minor qualification – relative to the intermediate case of "extended single profiles" – but as it stands, it is perhaps the quickest way of justifying the claim that an axiomatisation of utilitarianism be multi-profile. There is a further important, although not logically compelling argument that draws upon the intellectual history of the field. Be it for better or worse, Arrow's multi-profile approach has moulded social choice theory. Most of the relevant literature on utilitarian and related rules is now phrased in the related language of "social welfare functionals" (SWFL): d'Aspremont and Gevers (1977), Deschamps and Gevers (1978, 1979), Maskin (1978), Blackorby et al. (1982, 1984), Roberts (1980a). These theorems have established the customary standards that must be met for a logical derivation of utilitarianism to count as an axiomatisation. If only for that reason and for the related one of making comparisons easier, it is desirable to construct a SWFL equivalent of Harsanyi's Aggregation Theorem.

[4] Sen seems to reserve his point (2) for the Aggregation Theorem and his point (1) for the Impartial Observer Theorem. In each case Weymark (1991) extends Sen's critique to the other theorem.

A step in this direction was taken by Coulhon and Mongin (1989), who indeed constructed a multi-profile version of the theorem. But their definitions of a SWFL - as a function from the set of utility profiles into a *utility* set - departed from the standard one in social choice theory, and they were led to introduce a strong axiom of *cardinal* Independence of Irrelevant Alternatives. As a result of this unconventional framework, comparison between Coulhon and Mongin's multi-profile result and the above-mentioned characterizations of utilitarianism was somewhat complex. It could have been objected that cardinal Independence posed an exacting informational requirement on the social observer. In contrast to this early attempt the present paper offers a completely classical SWFL treatment of the Aggregation Theorem. The new multi-profile result turns out to be easily obtainable, after the very weak algebraic framework of *mixture sets* used throughout Coulhon and Mongin (1989) is strengthened. What is required here is that the set of prospects be a convex subset of a vector space. This is still a weak, easily acceptable algebraic assumption. It is satisfied in the banal case where prospects are taken to be probability distributions over a preexisting set of "prizes".

The paper is organised as follows. Section 2 states and proves the main result. Section 3 compares it with antecedent characterisations of utilitarianism as well as with the claim that a von Neumann-Morgenstern assumption posed on individual utilities does not provide a way of escape from Arrow's Impossibility Theorem. This comparison gives rise to two further results in social choice theory. Section 4 elaborates upon interpretative issues. Besides discussing the symmetry requirement, we address the more delicate issue of interpersonal comparisons of utility. The latter comes into the open once Propositions 1 and 2 are compared with the classic characterisation of utilitarianism by d'Aspremont and Gevers (1977). We end by concluding that *at least one* major obstacle to the use of the Aggregation Theorem as an argument for utilitarianism has been removed.

2. A SWFL version of the aggregation theorem

2.1. Algebraic preliminaries

We take the alternative set $\mathscr{M}$ to be a *convex subset of a vector space* $\mathscr{V}$. For any $x, y \in \mathscr{M}$ and any $\lambda \in [0,1]$ denote by $x\lambda y$ the convex combination, or mixture, $\lambda x+(1-\lambda)y$. We single out for special interest *mixture-preserving* (MP) functions. Formally, $u: : \mathscr{M} \to \mathbb{R}$ is MP if

$$\forall (x,y) \in \mathscr{M}^2 , \quad \forall \lambda \in [0,1], u(x\lambda y) = \lambda u(x) + (1-\lambda)u(y) \ .$$

A function $U: \mathscr{M} \to \mathbb{R}^n$ that is MP component by component will also be called MP. Denote by $\mathscr{L}(\mathscr{M})$ the set of all MP functions on $\mathscr{M}$. Clearly, it is a vector space and all the constant functions are in it. Also, note the following fact that we shall use repeatedly in this article:

Fact 1. *If $\mathscr{M}$ is a convex subset of a vector space $\mathscr{V}$, $u \in \mathscr{L}(\mathscr{M})$ if and only if u is affine on $\mathscr{M}$ (i.e., if and only if there is a linear form φ and a constant function k on $\mathscr{V}$ such that, for all $x \in \mathscr{M}$, $u(x) = \varphi(x) + k$).*

A proof of this useful lemma can be found, for instance, in Coulhon and Mongin (1989, p 183). They state the result for the case where $\mathscr{V} = \mathbb{R}^n$, but their argument can easily be extended to any linear space.

The assumption made here on the alternative set encompasses the usual one, according to which there is a measurable set $(P, \mathscr{A}(P))$ of results and the prospect set is defined to be $\Delta(P)$, the set of all probability distributions over $(P, \mathscr{A}(P))$. $\Delta(P)$ is known to be a convex subset of the vector space of all *signed* measures over $(P, \mathscr{A}(P))$ (see, e.g., Halmos, 1950, Chap. 6). (If P is finite, $\Delta(P)$ of course reduces to the unit simplex of some $\mathbb{R}^{\ell}$.) In this context, the notion of a MP function captures that of a von Neumann-Morgenstern (VNM) utility function. Our algebraic assumption is also applicable to a context in which there is no uncertainty, the feasible set is convex and utilities are affine; the same notation can serve in either context. Nevertheless, the assumption of this paper turns out to be slightly stronger than that of a *mixture set* in the sense considered in Fishburn (1982). The connection between mixture sets and convex subsets of vector spaces is studied by Mongin (1990), who provides the axioms required to go from the weaker to the stronger concept. As far as the social choice exercise is concerned, it is more appropriate to assume the convex structure right from the start.

The *dimension* of $\mathscr{M}$, denoted by $\dim\mathscr{M}$, is the dimension of the affine span of $\mathscr{M}$, denoted by $\mathrm{Aff}(\mathscr{M})$, that is the dimension of the *linear* span of the translated convex $\mathscr{M}-x_0$, for any $x_0 \in \mathscr{M}$; denote the latter span by $\mathrm{Vect}(\mathscr{M}-x_0)$. From this definition, $\dim\mathscr{M}=0$ if and only if $\mathscr{M}$ is a singleton and $\dim\mathscr{M}=1$ if and only if $\mathscr{M}$ is isomorphic to an interval of the real line. For $x_0, x_1, \dots, x_n \in \mathscr{V}$, define the family $\{x_0, x_1, \dots, x_n\}$ to be *affinely independent* if for any $j \in \{0, 1, \dots, n\}$, the family $\{x_i - x_j\}_{\substack{i=0,1,\dots,n\\ i\neq j}}$ is linearly independent. Note that $\{x_0, x_1, \dots, x_n\}$ is affinely independent if and only if there is no $j \in \{0, 1, \dots, n\}$, such that x_j can be written as an *affine* combination of the remaining x_i's, that is, there is no $(\lambda_i)_{i\neq j} \in \mathbb{R}^n$ with $\sum_{i\neq j} \lambda_i = 1$ such that

$$x_j = \sum_{i\neq j} \lambda_i x_i \ .$$

Henceforth, when we speak of an "independent" family, or of "independent" elements, we mean affine rather than linear independence. (The definitions in this paragraph are standard ones, e.g., Kelly 1979.)

The connection between the dimensionality of $\mathscr{M}$ and independence is as follows: $\dim\mathscr{M}=n$ if and only if there is a maximal independent family of $n+1$ elements in $\mathrm{Aff}(\mathscr{M})$. We can actually strengthen the connection in the following useful way:

Fact 2. *Suppose $\mathscr{M}$ is a convex subset of a vector space $\mathscr{V}$. Then $\dim\mathscr{M}=n$ if and only if there is a maximal independent family of $n+1$ elements in $\mathscr{M}$.*

Note also that Fact 1 provides an easy way of computing $\dim\mathscr{L}(\mathscr{M})$ when $\mathscr{M}$ is finite-dimensional: $\dim\mathscr{L}(\mathscr{M}) = \dim\mathscr{M}+1$.

Before turning to the social choice exercise it is useful to clarify the following algebraic problem. Take $\{x_0, \dots, x_n\}$ in $\mathscr{M}$ and let us see whether or not there is a function $u \in \mathscr{L}(\mathscr{M})$ such that $u(x_i)=a_i$, $i=0,\dots,n$, where $(a_0,\dots,a_n)$ is a $(n+1)$-tuple of predetermined values. If $\{x_0,\dots,x_n\}$ is affinely dependent it is obvious that u may not exist. (Suppose for instance that $x_0 = x_1 \lambda x_2$ for some $\lambda \in [0,1]$ and that $a_1 = a_2$, but $a_0 \neq a_1$.) Now if $\{x_0,\dots,x_n\}$ is affinely independent, the answer is in the affirmative. To elaborate somewhat upon this intuitively clear result, consider the family $\{x_1 - x_0, \dots, x_n - x_0\}$. We know that it is linearly independent. So we can define: $u'(x_i - x_0) = a_i - a_0$, $1 \leq i \leq n$ and extend u' into a linear form on $\mathscr{V}$, using a construction which is identical to that used in the proof of the Hahn-Banach Theorem (Zorn's Lemma, or some variant of it, will be needed). Now, define for all $\xi \in \mathscr{V}: u(\xi) = u'(\xi - x_0) + a_0$. The restriction of u to $\mathscr{M}$ has the desired properties and Fact 1 implies that it is in $\mathscr{L}(\mathscr{M})$. Formally:

Fact 3. *Suppose that $\{x_0, \ldots, x_n\}$ is an affinely independent family in $\mathscr{M}$. For any $(a_0, \ldots, a_n) \in \mathbb{R}^{n+1}$, there is $u \in \mathscr{L}(\mathscr{M})$ such that $u(x_i) = a_i$, $i = 0, \ldots, n$.*

The problem stated at the beginning of last paragraph may be referred to as that of *constrained attainability*, since the set $\{x_0, \ldots, x_n\}$ was given at the start. In contrast, the more trivial problem of *unconstrained attainability* requires one to find any $u \in \mathscr{L}(\mathscr{M})$ and $x_0, \ldots, x_n \in \mathscr{M}$ such that $u(x_i) = a_i$, $i = 0, \ldots, n$ for some predetermined vector $(a_0, \ldots, a_n)$. Fact 3 implies that the unconstrained attainability problem can be solved if $\dim \mathscr{M} \geq n$.

The social-choice-theoretic relevance of Fact 3 is plain. Characterisations of social choice rules in a multi-profile context have usually relied on Arrow's Unrestricted Domain assumption as applied to a set of profiles which are not subject to any algebraic restrictions. Hence, provided that there are enough (usually more than 2) distinct alternatives, it has been possible to assume at various stages in the proofs that there exist utility profiles with particular predetermined values. Here, we consider all functions in $\mathscr{L}(\mathscr{M})$, but no other ones; consequently, we cannot reproduce the usual mode of proof unless we check that our attainability assumptions are warranted by the algebraic structure. It will be seen that, for the proofs in the next section to carry through, we must make the benign assumption that $\dim \mathscr{M} \geq 2$.[5]

2.2. Definitions

Some social-choice-theoretic vocabulary must now be introduced. The number of individuals in society is $n \geq 2$.

Define a *social welfare functional* (SWFL) to be a function:

$$F: \mathscr{L}(\mathscr{M})^n \to 2^{\mathscr{M}^2}$$

such that for every $U = (u_1, \ldots, u_n) \in \mathscr{L}(\mathscr{M})^n$, $F(U)$ is a weak ordering (i.e., it is transitive, reflexive, and complete.) As usual, $F(U)$ describes the observer's weak preference on the set of alternatives $\mathscr{M}$. Denote by $I(U)$ and $P(U)$ the symmetric and asymmetric parts of $F(U)$, respectively. We shall impose the following von Neumann-Morgenstern restrictions on F:

---

[5] Explicit attainability conditions often occur in the single-profile theory of social choice, especially in the *extended* single-profile approach explored by K. Roberts (1980b). They were used in most of the early proofs of the Aggregation Theorem but turned out to be dispensable on further examination (Coulhon and Mongin, 1989, p 184). It remains to fully clarify the logical role played by attainability conditions in the *subjective uncertainty* analogues of the theorem – see note 2. Mongin's Savage-based variant does not require any such condition, but Broome's Jeffrey-based version appears to need one.

(VNM) = (VNM1) & (VNM2)

(VNM1) For all $U \in \mathscr{L}(\mathscr{M})^n$, $F(U)$ is continuous in the following sense:

$\forall (x, y, z) \in \mathscr{M}^3$

$\{\lambda \in [0,1] : zF(U)(x\lambda y)\}$ and $\{\mu \in [0,1] : (x\mu y)F(U)z\}$

are closed subsets of $[0,1]$;

(VNM2) $F(U)$ satisfies independence, that is:

$\forall (x, y, z) \in \mathscr{M}^3, \forall \lambda \in]0,1]$

$xF(U)y \Leftrightarrow (x\lambda z)F(U)(y\lambda z)$.

This is but standard expected utility theory as applied to social orderings $F(U)$; we have selected a non-independent but convenient axiomatisation.[6] (VNM1) and (VNM2) will sometimes be used separately from each other in the sequel because only the latter turns out to be crucial to utilitarianism. Define the following conditions:

Independence of irrelevant alternatives:

(I) $\forall U \in \mathscr{L}(\mathscr{M})^n, \forall U' \in \mathscr{L}(\mathscr{M})^n, \quad \forall (x, y) \in \mathscr{M}^2,$

$$\left.\begin{array}{l} U(x) = U'(x) \\ U(y) = U'(y) \end{array}\right\} \Rightarrow xF(U)y \quad \text{iff} \quad xF(U')y \ ;$$

Pareto Indifference:

(PI) $\forall U \in \mathscr{L}(\mathscr{M})^n, \quad \forall (x, y) \in \mathscr{M}^2,$

$U(x) = U(y) \Rightarrow xI(U)y$;

Strong Pareto:

(SP) = (SP1) & (SP2).

(SP1) $\forall U = (u_1, \ldots, u_n) \in \mathscr{L}(\mathscr{M}), \quad \forall (x, y) \in \mathscr{M}^2$

$u_i(x) \geq u_i(y), \quad i = 1, \ldots, n$

[denoted by $U(x) \geq U(y)$]

$\Rightarrow xF(U)y$;

[6] Compare Fishburn, 1970. The particular version of von Neumann and Morgenstern's independence that we are using here resembles Samuelson's in the early days of expected-utility theory (see Fishburn 1989, for the list of available variants).

(SP2) $\forall U=(u_1,\ldots,u_n)\in\mathscr{L}(\mathscr{M})\ ,\quad \forall(x,y)\in\mathscr{M}^2\ ,$
$U(x)\geq U(y)\ ,\quad$ and $\exists j:1\leq j\leq n, u_j(x)>u_j(y)$
[denoted by $U(x)>U(y)$]
$\Rightarrow xP(U)y\ .$

Clearly (SP1) $\Rightarrow$ (PI).

2.3. The result

Let us now show that the Pareto conditions, Independence of Irrelevant Alternatives, and the von Neumann-Morgenstern restriction imply that the social choice rule can be represented by an affine combination of the individuals' utilities.

Proposition 1. *Suppose that* $\dim\mathscr{M}\geq 2$ *and* F *satisfies* (I), (PI) *and* (VNM). *Then there is a vector* $(a_1,\ldots,a_n)\in\mathbb{R}^n$, *unique up to a positive scale factor, such that for all* $U=(u_1,\ldots,u_n)\in\mathscr{L}(\mathscr{M})^n$, $F(U)$ *can be represented by* $\sum a_iu_i$, *i.e.,*

$$\forall x,y\in\mathscr{M}, xF(U)y \quad\text{iff}\quad \sum a_iu_i(x)\geq\sum a_iu_i(y)\ .$$

If (PI) *is strengthened into* (SP1), *the* a_i *are nonnegative; if it is strengthened into* (SP), *the* a_i *are strictly positive.*

The proof goes as follows.

Lemma 1.

$$\forall U,U'\in\mathscr{L}(\mathscr{M})^n\ ,\quad \forall x,y,x',y'\in\mathscr{M}\ ,$$

$$\left.\begin{array}{l}U(x)=U'(x')\\ U(y)=U'(y')\end{array}\right\}\Rightarrow xF(U)y \quad\text{iff}\quad x'F(U')y'\ .$$

Proof. Let us first prove the following particular case of Lemma 1:

$$\forall U,U'\in\mathscr{L}(\mathscr{M})^n\ ,\quad \forall x,y,x'\in\mathscr{M}\ ,$$

$$\left.\begin{array}{l}U(x)=U'(x')\\ U(y)=U'(y)\end{array}\right\}$$

$$\Rightarrow xF(U)y \quad\text{iff}\quad x'F(U')y\ ,\qquad\text{and}\quad yF(U)x \quad\text{iff}\quad yF(U')x'\ .$$

Suppose that $\{x,y,x'\}$ is independent. Then, from Fact 3 above there is $U''\in\mathscr{L}(\mathscr{M})$ as in the table:

	x	y	x'
U	$U(x)$	$U(y)$	
U'		$U(y)$	$U(x)$
U''	$U(x)$	$U(y)$	$U(x)$

(I), (PI) and the transitivity of social preference imply that:

$xF(U)y$ iff $xF(U'')y$ iff $x'F(U'')y$ iff $x'F(U')y$, and
$yF(U)x$ iff $yF(U'')x$ iff $yF(U'')x'$ iff $yF(U')x'$.

Suppose now that $\{x, y, x'\}$ is dependent. From the condition $\dim \mathcal{M} \geq 2$ and Fact 2, it is easily seen that, say, $\{x, y\}$ can be completed into an independent set $\{x, y, w\}$ for some $w \in \mathcal{M}$. Now, $\{x', y, w\}$ is independent too, since $w \notin \mathrm{Aff}(\{x, y\}) = \mathrm{Aff}(\{x', y\})$. Fact 3 ensures that there are U^1, $U^2 \in \mathcal{L}(\mathcal{M})^n$ as in the following table:

	x	y	x'	w
U	$U(x)$	$U(y)$		
U'		$U(y)$	$U(x)$	
U^1	$U(x)$	$U(y)$		$U(x)$
U^2		$U(y)$	$U(x)$	$U(x)$

The conclusions that $xF(U)y$ iff $x'F(U')y$ and $yF(U)x$ iff $yF(U')x'$ again follow from alternating the application of (I) and (PI), as well as using the transitivity of social preference.

The result that has just been proved extends straightforwardly to the general case where $y \neq y'$. Using Fact 3 again, we can find $U''' \in \mathcal{L}(\mathcal{M})^n$ as in the table below, and then apply this result:

	x	y	x'	y'
U	$U(x)$	$U(y)$		
U'			$U(x)$	$U(y)$
U'''		$U(y)$	$U(x)$	

Lemma 1 is a welfarism theorem (see d'Aspremont 1985, or Sen 1986). The proof differs from the usual one because of the requirement that all the auxiliary functions be in $\mathcal{L}(\mathcal{M})$. Fact 3 and the assumption that $\dim \mathcal{M} \geq 2$ were used to solve constrained attainability problems. In the proofs below unconstrained attainability problems occur repeatedly and are solved by the same argument without our mentioning it explicitly.

Lemma 2. *The relation R on $\mathbb{R}^n$ defined by:*

$$\forall (a,b) \in \mathbb{R}^n \times \mathbb{R}^n \; aRb \quad \text{iff}$$

$$\exists\, U \in \mathcal{L}(\mathcal{M})^n, (x,y) \in \mathcal{M}^2 : U(x) = a, U(y) = b, xF(U)y$$

is an ordering. Moreover, it is continuous in the (VNM1) *sense:*

(VNM1*) *for any a_1, a_2, a_3 in $\mathbb{R}^n$, $\{\lambda \in [0,1] : a_3 R(a_1 \lambda a_2)\}$ and $\{\mu \in [0,1] : (a_1 \mu a_2) R a_3\}$ are closed subsets of $[0,1]$;*
and it satisfies independence in the (VNM2) *sense:*

(VNM2*) *for any a_1, a_2, a_3 in $\mathbb{R}^n$, any $\lambda \in \,]0,1]$, $a_1 R a_2 \Leftrightarrow (a_1 \lambda a_3) R (a_2 \lambda a_3)$.*

Proof. Reflexivity of R follows from the fact that any $a \in \mathbb{R}^n$ belongs to the range of a $U \in \mathcal{L}(\mathcal{M})^n$, along with the fact that $F(U)$ itself is reflexive.

Completeness of R follows from the fact that any two points, a, b in $\mathbb{R}^n$ belong to the range of some $U \in \mathcal{L}(\mathcal{M})^n$, along with the fact that $F(U)$ itself is complete.

To establish transitivity, take $a, b, c \in \mathbb{R}^n$ such that aRb and bRc. Hence, there are $U, U' \in \mathscr{L}(\mathscr{M})$, and x, y, x', y' such that

$$U(x)=a\ , \qquad U(y)=b\ , \qquad xF(U)\,y$$
$$U'(x')=b\ , \qquad U'(y')=c\ , \qquad x'F(U')y'\ .$$

We can find $U''' \in \mathscr{L}(\mathscr{M})^n$ such that $U'''(x_1)=a$, $U'''(x_2)=b$, and $U'''(x_3)=c$ for some $\{x_1, x_2, x_3\}$ independent in $\mathscr{M}$. Lemma 1 implies that $x_1 F(U''')\,x_2$ and $x_2 F(U''')\,x_3$. So transitivity of $F(U''')$ ensures that $x_1 F(U''')\,x_3$ and aRc.

Now, given a_1, a_2, a_3 in $\mathbb{R}^n$, we wish to show that $\{\lambda \in [0,1]: a_3 R(a_1 \lambda a_2)\}$ and $\{\mu \in [0,1]: (a_1 \mu a_2)\,Ra_3\}$ are closed subsets of $[0,1]$. There exists $U \in \mathscr{L}(\mathscr{M})^n$ such that $U(x_1)=a_1$, $U(x_2)=a_2$ and $U(x_3)=a_3$ for some x_1, x_2, x_3 in $\mathscr{M}$. Hence, using the MP property of U we see that we have to prove the closedness of:

$$\{\lambda \in [0,1]: U(x_3)\,RU(x_1 \lambda x_2)\} \quad \text{and}$$
$$\{\mu \in [0,1]: U(x_1 \mu x_2)\,RU(x_3)\}\ .$$

This will result from (VNM1) if we prove that:

$$\{\lambda \in [0,1]: U(x_3)\,RU(x_1 \lambda x_2)\}=\{\lambda \in [0,1]: x_3 F(U)(x_1 \lambda x_2)\} \tag{1}$$

and

$$\{\mu \in [0,1]: U(x_1 \mu x_2)\,RU(x_3)\}=\{\mu \in [0,1]: (x_1 \mu x_2)\,F(U)\,x_3\}\ . \tag{2}$$

The inclusion from right to left in equation (1) is trivial. To check that the converse inclusion also holds, take any λ such that $U(x_3)\,RU(x_1 \lambda x_2)$. There are $U' \in \mathscr{L}(\mathscr{M})^n$ and $x, y \in \mathscr{M}$ satisfying the properties that $U'(x)=U(x_3)$, $U'(y)=U(x_1 \lambda x_2)$, $xF(U')y$; hence (from Lemma 1) $x_3 F(U)(x_1 \lambda x_2)$, as requested. A parallel argument takes care of (2).

Finally, given a_1, a_2, a_3 in $\mathbb{R}^n$ and any $\lambda \in\,]0,1]$, we wish to show that

$$a_1 Ra_2 \Leftrightarrow (a_1 \lambda a_3)\,R(a_2 \lambda a_3)\ .$$

The proof consists in taking U as in the previous paragraph, applying its MP property, (VNM2), as well as Lemma 1. Details are left for the reader. □

Lemma 3. *The ordering R defined in Lemma 2 satisfies:*

(SP1*) *for all $a, b \in \mathbb{R}^n$, $a \geq b \Rightarrow aRb$,*
whenever (SP1) *holds, and:*

(SP2*) *for all $a, b \in \mathbb{R}^n$, $a > b \Rightarrow aRb$ and not bRa,*
whenever (SP2) *holds.*

Proof. Given $a, b \in \mathbb{R}^n$, there are $U \in \mathscr{L}(\mathscr{M})^n$, $x, y \in \mathscr{M}$ such that $U(x)=a$ and $U(y)=b$. Hence, if $a \geq b$, $xF(U)\,y$ follows from (SP1), and aRb holds. Similarly, if $a > b$, $xP(U)\,y$ follows from (SP2), and aRb holds. Assume that bRa also holds. Then, there are $U' \in \mathscr{L}(\mathscr{M})^n$, $x', y' \in \mathscr{M}$ such that $U'(x')=a$, $U'(y')=b$ and $y'F(U')\,x'$; Lemma 1 implies that $yF(U)\,x$, a contradiction. □

End of the proof of Proposition 1:

From Lemmas 1 and 2, we know that F has given rise to a (uniquely defined) ordering R on $\mathbb{R}^n$ which satisfies (VNM1) and (VNM2). Applying the expected utility theorem (e.g., Fishburn 1982, Chap. 1), we conclude that R can be represented by a function $G: \mathbb{R}^n \to \mathbb{R}$ which is MP on $\mathbb{R}^n$, and that the set of MP representations of R is exactly the invariance class of G with respect to positive affine transformations. Now, Fact 1 states that mixture preservation of $\mathbb{R}^n$ is equivalent to the property of being affine on $\mathbb{R}^n$. Hence, there are real numbers $a_1, \ldots, a_n, b$ such that

$$\forall X = (x_1, \ldots, x_n) \in \mathbb{R}^n \ , \qquad G(X) = \sum a_i x_i + b \ .$$

Any alternative choice of numbers $a_1', \ldots, a_n'$, b' such that $\sum a_i' x_i + b'$ represents R must be such that $(a_1', \ldots, a_n') = \alpha\,(a_1, \ldots, a_n)$ for some $\alpha > 0$. When stronger Pareto conditions than (PI) hold, sign restrictions on the a_i follow from applying Lemma 3 to suitably chosen $X \in \mathbb{R}^n$. □

The conclusion of the *multi-profile* Aggregation Theorem involves weights a_i that are not only independent of the particular profile U but also essentially unique. The weights mentioned in the conclusion of the *single-profile* theorem are not independent; they may or may not be unique, depending on whether or not the u_i in the given profile $U = (u_1, \ldots, u_n)$ are affinely independent (Fishburn 1984; Coulhon and Mongin 1989).

Another word of comparison relates to the sign of the coefficients. Similar results to those of Proposition 1 have already been reached in the single-profile context. Relying as it does on Pareto-indifference only, the basic version of the Aggregation Theorem cannot ensure any sign restriction on the single-profile coefficients. Further work by Fishburn (1984) and Weymark (1993, 1994) has shown that stronger Pareto conditions entail that the single-profile coefficients can be chosen to be nonnegative. They can be chosen to be strictly positive in the case of a condition analogous to (SP) here.[7]

For the sake of later discussion we record the fact that the completeness assumption posed on the observer's preference plays a limited role in the above derivations. It is used at the obvious place in the proof of Lemma 2 as well as – implicitly – throughout the end of the proof. It is unnecessary to prove Lemma 1 and Lemma 3.

3. A comparison of theorems on utilitarianism

Once restated in the language of contemporary social choice theory, the Aggregation Theorem appears somewhat atypical. Sen's (1970, 1974) pioneering work on the informational bases of welfare judgements resulted in the discovery that Arrow's theorem can be transformed into a variety of possibility results. In the same vein, the writers on utilitarianism in the 70's have derived this rule from suitably chosen *invariance principles*; the latter formalise the possibility of cardinal comparisons of utility some way or another. In contrast, the derivation of generalised utilitarianism in the last section does not apparently exploit any invariance principle. Because invariance is a consequence of the utilitarian rule itself, it must be hidden somewhere in the axiom set. The task of this section is to clarify the formal connection between the VNM hypothesis and invariance principles à la Sen, while the next one explores the corresponding interpretative

[7] For a summary and quick proof of these results, see De Meyer and Mongin (1994). Notice that it is more difficult for the single-profile approach to apply the stronger Pareto conditions and sign the coefficients than it is for the present one. The reason for this discrepancy is the non-uniqueness of single-profile coefficients in case of affinely dependent utilities.

issue. As it turns out, clarifying the formal connection has two interesting by-products: a variant of Proposition 1 which disposes of the continuity assumption (Proposition 2) and a variant of Arrow's theorem adapted for a domain of VNM utilities (Proposition 3).

We shall state invariance postulates with reference to any choice set X of at least three elements. The set $\mathscr{D}$ of admissible utility profiles $U=(u_1,\ldots,u_n)$ $(n\geq 2)$ is left unspecified at this stage. A *social welfare functional* is now any

$$F:\mathscr{D}\to 2^{X^2}$$

where $F(U)$ is a weak ordering. For any $\beta\in\mathbb{R}$, $\hat{\beta}$ will denote the vector $(\beta,\ldots,\beta)$ in $\mathbb{R}^n$. Define

Difference comparability:

(DC) $\forall U\in\mathscr{D}, \quad \forall\beta\in\mathbb{R}$ s.t. $U+\hat{\beta}\in\mathscr{D}$,

$$F(U)=F(U+\hat{\beta})\ ;$$

Cardinality and unit comparability:

(CU) $\forall U\in\mathscr{D}, \quad \forall\alpha>0, \quad \forall\boldsymbol{\beta}\in\mathbb{R}^n$ s.t. $\alpha U+\boldsymbol{\beta}\in\mathscr{D}$,

$$F(U)=F(\alpha U+\boldsymbol{\beta})\ ;$$

Cardinality and full comparability:

(CC) $\forall U\in\mathscr{D}, \quad \forall\alpha>0, \quad \forall\beta\in\mathbb{R}$ s.t. $\alpha U+\hat{\beta}\in\mathscr{D}$,

$$F(U)=F(\alpha U+\hat{\beta})\ .$$

The existing characterisations of utilitarianism rely on one of these invariance principles; cf. Roberts (1980a): (DC); d'Aspremont and Gevers (1977): (CU) [denoted by (IOU) in their paper]; Maskin (1978), as well as Deschamps and Gevers (1978, 1979): (CC); Blackorby and al. (1982, 1984): (CU) or (CC).

Assuming ***Unrestricted Domain*** (UD), i.e., that $\mathscr{D}=(\mathbb{R}^n)^X$, welfarism follows from (I) and (PI), i.e., F is (uniquely) associated with an ordering R on $\mathbb{R}^n$ which is defined as in Lemma 2 of Sect. 2. (As usual, denote by P and I the strict preference and indifference relations on $\mathbb{R}^n$ associated with R.) Corresponding invariance properties follow from (DC), (CU) and (CC):

(DC*) $\forall\bar{U},\bar{V}\in\mathbb{R}^n,\forall\hat{\beta}=(\beta,\ldots,\beta)\in\mathbb{R}^n$,

$$\bar{U}R\bar{V} \text{ iff } (\bar{U}+\hat{\beta})R(\bar{V}+\hat{\beta})\ ;$$

(CU*) $\forall\bar{U},\bar{V}\in\mathbb{R}^n,\forall\alpha>0,\forall\boldsymbol{\beta}\in\mathbb{R}^n$,

$$\bar{U}R\bar{V} \text{ iff } (\alpha\bar{U}+\boldsymbol{\beta})R(\alpha\bar{V}+\boldsymbol{\beta})\ ;$$

(CC*) $\forall\bar{U},\bar{V}\in\mathbb{R}^n,\forall\alpha>0,\forall\hat{\beta}=(\beta,\ldots,\beta)\in\mathbb{R}^n$,

$$\bar{U}R\bar{V} \text{ iff } (\alpha\bar{U}+\hat{\beta})R(\alpha\bar{V}+\hat{\beta})\ .$$

Analogous VNM and Pareto properties, i.e., (VNM1*), (VNM2*), (SP1*), (SP2*), have already been defined in the course of proving Proposition 1. We are now ready to compare most of the available characterisations of utilitarianism, including the above restatement of the Aggregation Theorem, within a single formal framework.

We note first that (CU*) implies (VNM2*). To see that the converse holds, take $\bar{U}, \bar{V} \in \mathbb{R}^n$. Then, for any α in $]0,1]$,

$$\bar{U} R \bar{V} \quad \text{iff} \quad (\alpha \bar{U}) R (\alpha \bar{V})$$

follows from applying (VNM2*) to the mixtures $\bar{U}\alpha 0$ and $\bar{V}\alpha 0$. The case $\alpha > 1$ is easily disposed of, and, for any $\bar{W} \in \mathbb{R}^n$,

$$\bar{U} R \bar{V} \quad \text{iff} \quad (\bar{U} + \bar{W}) R (\bar{V} + \bar{W})$$

follows from considering the mixtures $\bar{U}\frac{1}{2}\bar{W}$ and $\bar{V}\frac{1}{2}\bar{W}$. Let us formally state this as:

Lemma 4. (VNM2*)$\Leftrightarrow$(CU*).

This simple observation answers the question raised above: Once welfarism is granted, it is equivalent to assume that the social observer satisfies the independence part of the (VNM) axioms, or that he is endowed with a (CU) informational basis. The curious social-choice-theoretic implications of this equivalence are pursued in Sect. 4.

Comparison of Lemma 4 with the invariance-based characterisations of utilitarianism suggests that a variant of Proposition 1 could be constructed by dropping the continuity axiom (VNM1) imposed on the observer's preference. However, one should expect the weaker asumption set to imply a weaker utilitarian result. Some further terminology is useful at this juncture. After d'Aspremont (1985, p 46) we shall define *generalised utilitarianism* as follows: there are non-negative numbers $\alpha_1, \ldots, \alpha_n$, one of which is strictly positive, such that for all $U = (u_1, \ldots, u_n) \in \mathscr{D}$ and for all $x, y \in X$,

(GU) $\quad \sum \alpha_i u_i(x) \geq \sum \alpha_i u_i(y) \Leftrightarrow x F(U) y$.

Weak utilitarianism is defined by replacing (GU) with the condition that

(WU) $\quad \sum \alpha_i u_i(x) > \sum \alpha_i u_i(y) \Rightarrow x P(U) y$.

In the presence of the completeness assumption made on the observer's preference, (WU) implies that:

$$x P(U) y \Rightarrow \sum \alpha_i u_i(x) \geq \sum \alpha_i u_i(y) \,, \qquad \text{and}$$
$$x I(U) y \Rightarrow \sum \alpha_i u_i(x) = \sum \alpha_i u_i(y) .$$

To go from this seemingly unusual rule to generalised utilitarianism, one and only one piece of information should be added:

$$\sum \alpha_i u_i(x) = \sum \alpha_i u_i(y) \Rightarrow x I(U) y \,.$$

That is, weak utilitarianism is compatible with the fact that hyperplanes $\sum \alpha_i X_i = k$ in the utility space have points which are not indifferent to each other (in the sense of the induced I on $\mathbb{R}^n$). Note also that both rules include *dictatorship* as a particular case. As usual, the latter is defined by the property that there is an individual i such that for all $U \in \mathscr{D}$ and all $x, y \in X$,

(D) $\quad u_i(x) > u_i(y) \Rightarrow x P(U) y$.

In terms of the definitions just introduced, Proposition 1 derived a strengthened (GU) rule – in which all individual coefficients are positive – from the domain assumption, (I), (SP) and (VNM). Proposition 2 below derives (WU) from the weaker axiom set in which (VNM) is replaced with its (VNM2) part only. We state it in such a way that it also encompasses a related result in the invariance-based approach to utilitarianism.

Proposition 2. *Suppose that either (i) X is any set with $\# X \geq 3$ and $\mathscr{D} = (\mathbb{R}^n)^X$ or (ii) X is a convex subset $\mathscr{M}$ of a vector space with* $\dim \mathscr{M} \geq 2$ *and $\mathscr{D} = \mathscr{L}(\mathscr{M})^n$. Assume* (I), (SP) *and either* (CU) *if (i) or* (VNM2) *if (ii). Then weak utilitarianism holds.*

Proof. The assumptions in case (i) imply welfarism as well as (CU*) and (SP*) = (SP$_1^*$) & (SP$_2^*$). Lemmas 1, 2, 3 in Sect. 2 have shown that the assumptions in case (ii) imply welfarism as well as (VNM2*) and (SP*). From Lemma 4 we can deal with the two cases at the same time by assuming an ordering R on $\mathbb{R}^n$, (CU*) and (SP*). Using (CU*) and transitivity of R, it is easy to see that $S_0 = \{a \mid a \in \mathbb{R}^n \text{ and } aR0\}$ is convex. (Take $a, b \in S_0$ and $\lambda \in [0, 1]$; hence $\lambda a R 0$, $(1-\lambda) b R 0$, $\lambda a + (1-\lambda) b R (1-\lambda) b$ and $\lambda a + (1-\lambda) b R 0$.) It is also the case that $T_0 = \{a \mid a \in \mathbb{R}^n \text{ and } 0 P a\}$ is convex. Note that S_0 is nonempty because R is reflexive and (SP2*) implies that T_0 is not empty. Of course $S_0 \cap T_0 = \emptyset$, and a standard separating hyperplane theorem ensures that there are real numbers $\alpha_1, \ldots, \alpha_n$, not all of them zero, and α, such that for any $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$,

$$a \in S_0 \Rightarrow \sum \alpha_i a_i \geq \alpha$$

$$a \in T_0 \Rightarrow \sum \alpha_i a_i \leq \alpha \ .$$

Using (SP2*) and standard limiting arguments, we have $\alpha_i \geq 0$, $i = 1, \ldots, n$, and $\alpha = 0$. Now, if we prove that for any $a \in \mathbb{R}^n$,

$$\sum \alpha_i a_i > 0 \Rightarrow a P 0 \ ,$$

(WU) will follow from (CU*), the definition of P and the welfarism lemma. Take a such that $\sum \alpha_i a_i > 0$. Then, $a \notin T_0$, whence $a \in S_0$, since R is complete. Clearly, it follows from (SP*) that there is an $a' \in S_0$ such that $a P a'$. If one had $a I 0$, one would have $0 P a'$ and $a' \in T_0$, a contradiction. □

The statement relative to (i) in Proposition 2 is Theorem 3.3.3 in d'Aspremont (1985, p 48–50) and is closely related to Theorem 2 in Gevers (1979). Note that the proof given here is very simple, relying as it does on a basic convexity argument. It is in the style of earlier proofs by Blackwell and Girschick (1954) in decision theory and by Roberts (1980a) in social choice theory. To obtain an "if and only if" utilitarian rule, a *continuity* assumption is clearly needed, that is, the sets S_0 and T_0 in the proof should be not only convex but also closed and open, respectively. Accordingly, define ***Continuity in the Utility Space*** as:

(C*) $\quad \forall a \in \mathbb{R}^n, \{a' \mid a' \in \mathbb{R}^n \text{ and } a' R a\}$ and $\{a' \mid a' \in \mathbb{R}^n \text{ and } a R a'\}$ are closed in $\mathbb{R}^n$.

This is Maskin's (1978) notion of continuity. It is somewhat ad hoc since it is posed on $\mathbb{R}^n$ directly rather than derived from antecedent properties of X and F. Clearly, X does not have to be a topological space. Where $X = \mathcal{M}$ and $\mathcal{D} = \mathcal{L}(\mathcal{M})^n$ there is a way of circumventing the problem, which is to assume (VNM1); this leads back to Proposition 1 of last section.

Alternatively, one may introduce ***Anonymity:***

(A) $\forall\, U = (u_1, \ldots, u_n) \in \mathcal{D}, \forall\, \sigma$ permutation of $\{1, \ldots, n\}$,

$$F(U) = F(u_{\sigma 1}, \ldots, u_{\sigma n})$$

as well as the induced condition on $\mathbb{R}^n$:

(A*) $\forall\, \bar{U} = (\bar{u}_1, \ldots, \bar{u}_n) \in \mathbb{R}^n, \forall\, \sigma$ permutation of $\{1, \ldots, n\}$,

$$\bar{U} I (\bar{u}_{\sigma 1}, \ldots, \bar{u}_{\sigma n}) \ .$$

As is well-known (e.g., Blackorby, Donaldson, Weymark 1984, p 351–352, or d'Aspremont 1985), when welfarism is granted, (A*) and (CU*) imply (C*). Not only is the resulting rule an "if and only if" one, as in generalised utilitarianism, but it of course has equal weights. This rule is *standard utilitarianism:*

$$\forall\, U = (u_1, \ldots, u_n) \in \mathcal{D} \quad \text{and} \quad \forall\, x, y \in X \ ,$$

$$x F(U) y \Leftrightarrow \sum u_i(x) \geq \sum u_i(y) \ ,$$

and the characterisation just obtained is the classic one by d'Aspremont and Gevers (1977): Standard utilitarianism is equivalent to (UD), (I), (SP), (CU), and (A). Note that the proof of Proposition 2 makes it clear that *most* of the result of d'Aspremont and Gevers (i.e., what has been called here weak utilitarianism) is already contained in (UD), (I), (SP), and (CU). By the same token, most of our result is contained in the domain assumption, (I), (SP), and (VNM2). Anonymity is conceptually necessary to restore the classical utilitarian formula but it has a limited mathematical role.

A consequence of adding either continuity or anonymity is that dictatorship becomes impossible. This is obvious enough as far as anonymity goes. To see the related point with continuity, we may go back to Proposition 2 and assume either (C*) in case (i) or (VNM1*) in case (ii). Reworking the separation argument it can be shown that either case delivers (GU) with *positive* coefficients α_i. To further illustrate the effect of continuity we may note that the initial axiom set, i.e. (I), (SP), and either (CU) or (VNM2), was compatible with the rule that Gevers (1979, p 78) calls *lexical individual dictatorship.* In the case in which $n = 2$ this rule is defined as follows: there is a reindexing of individuals such that, using the new indexes, for any $U \in \mathcal{D}$ and any two $x, y \in X, x P(U) y$ if either $u_1(x) > u_1(y)$ or $u_1(x) = u_1(y)$ and $u_2(x) > u_2(y)$; and $x I(U) y$ if $u_i(x) = u_i(y)$, $i = 1, 2$. Clearly, the rule is devised to meet all of the Pareto conditions; it trivially satisfies (CU) and can be seen to satisfy (I). Being highly discontinuous, it does not survive once either (C*) or (VNM1*) is assumed.

Another point should be made in connection with invariance principles. There appears to be no agreement on which of the three candidates, (DC), (CU), and (CC) is most expressive of the utilitarian mode of making interpersonal comparisons. In particular, Maskin (1978) has taken d'Aspremont and Gevers to task for choosing (CU) instead of (CC). Unlike the former, the latter principle does not discriminate between the informational roles of rescaling and change in the origin. This is the sense in which (CC) may seem less arbitrary than (CU). There is, however, an argument for preferring (CU) over any other candidate. To the best of our knowledge, the list of invariance principles which are relevant to social choice theory is by now essentially closed and agreed upon. To be specific, let us refer to the list $\mathscr{T}$ of the principles surveyed in d'Aspremont (1985). Now, we may note that (CU) is the logically strongest invariance principle in $\mathscr{T}$ among those which are *implied* by generalised utilitarianism. Even more interestingly, when (CU) is used as a *partially sufficient* condition for generalised utilitarianism - along with standard social-choice-theoretic conditions -, it cannot be replaced with any stronger principle in $\mathscr{T}$; for the strengthening would result in a contradiction. This fact results from a well-known impossibility theorem by Sen. The principle which is next in $\mathscr{T}$ to (CU) is ***Cardinality and Noncomparability:***

$$\text{(CN)} \quad \forall U \in \mathscr{D}, \forall \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{R}^{+*})^n\,, \quad \forall \boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$$

$$\text{s.t. } (\alpha_1 u_1 + \beta_1, \ldots, \alpha_n u_n + \beta_n) \in \mathscr{D}\,,$$

$$F(U) = F(\alpha_1 u_1 + \beta_1, \ldots, \alpha_n u_n + \beta_n)\,.$$

Sen's result (1970, Theorem 8*2) implies that there is no nondictatorial F satisfying (UD), (I), (SP), and (CN).[8] Hence, (CU) is a borderline axiom in a certain formal sense. This feature may be enough to justify according it a special status in utilitarian theory, despite Maskin's comment that it is arbitrary.

The reference to Sen's impossibility theorem raises a final question for discussion: Is there a corresponding impossibility result in the von Neumann-Morgenstern theory of social choice pursued in this paper? There is indeed; for the welfarism lemma of Sect. 2.3 can easily be turned into an impossibility theorem.

Proposition 3. *Suppose $X = \mathscr{M}$, a convex subset of a vector space with* $\dim \mathscr{M} \geq 2$, *and $\mathscr{D} = \mathscr{L}(\mathscr{M})^n$. Any F satisfying* (I), (SP), *and* (CN) *must satisfy* (D).

Proof. Take any F satisfying (I), (SP), and (CN). From Lemma 1, (I) and (SP) imply welfarism. (CN) implies (CU), which implies (CU*) if welfarism holds, that is (VNM2*) from Lemma 4. Hence, by the proof of Proposition 2, weak utilitarianism holds. Combining the latter with the definition of (CN), we reach the following conclusion: There are nonnegative numbers $\alpha_1, \ldots, \alpha_n$, not all of them zero, such that for all $U \in \mathscr{L}(\mathscr{M})^n$, for all $x, y \in \mathscr{M}$, and for all $(a_1, \ldots, a_n) \in (\mathbb{R}^{+*})^n$,

$$(1) \quad \sum \alpha_i u_i(x) > \sum \alpha_i u_i(y)$$

implies that

$$(2) \quad \sum \alpha_i a_i u_i(x) \geq \sum \alpha_i a_i u_i(y)\,.$$

Now, assume that more than one i have positive coefficients, e.g., $\alpha_1 > 0$ and $\alpha_i > 0$ for some $i > 1$. The implication can be shown to be a contradiction as follows. From Fact 3 and the condition that $\dim \mathscr{M} \geq 2$ we can find $u_i \in \mathscr{L}(\mathscr{M})$, $i = 1, \ldots, n$ and $x, y \in \mathscr{M}$ such that $u_1(x) > u_1(y)$, $u_i(x) < u_i(y)$, $i = 2, \ldots, n$, and (1) holds. A suitable choice of a_i then falsifies (2). □

---

[8] Sen's exact statement involves the Weak Pareto condition and is therefore even more closely related to the current version of Arrow's Theorem than are the statement in the text and Proposition 3 below.

This proposition clarifies the sense in which it can be said that "the von Neumann-Morgenstern hypothesis" does not circumvent Arrow's Impossibility Theorem.[9] For it says that if the individuals have VNM preferences, and the observer will take no more into account than the ordinal properties of these VNM preferences, then there is no consistent social choice rule under the usual conditions (I) and (SP). To remove the impossibility, it is enough either to *minimally* weaken (CN) into (CU), or to *maximally* weaken (CN) into the lack of any explicit invariance principle while *adding* the crucial (VNM2) restriction that the observer too has VNM preferences.

Alternatively, Proposition 3 can be interpreted as a variant of Arrow's Theorem proved on a special *economic domain.* Recall from Sect. 2.1 that our algebraic framework encompasses the particular application of consumer theory in which the agents' utility functions are affine on a convex feasible set.

4. Conceptual issues solved and unsolved

This section returns to interpretative issues, drawing upon the better understanding of the Aggregation Theorem which has hopefully been gained from Propositions 1, 2, and 3. Our strategy here is to move from relatively unproblematic issues to more difficult ones, some of which will be left unanswered. *Anonymity* is perhaps the easiest of all. In the more standard, single-profile approach to the Aggregation Theorem, a weighted sum formula is first derived, then symmetric formulas – typically, either Harsanyi's average utility rule or the standard Benthamite sum rule – are obtained using one of the following procedures: (i) the imposition of a symmetry requirement on the social welfare function $W(u_1,\ldots,u_n)$; (ii) a suitable rescaling of the individuals' utilities. Now, (i) is simply not correct. The courteous way of showing this is to remark that the social welfare *function* is not a primitive term of the single-profile approach to the Aggregation Theorem; rather it is derived from the (single-profile) Pareto Indifference condition. Coming as it does in the middle of the proof, the symmetry assumption:

$$(*) \qquad W(u_1,\ldots,u_n)=W(u_{\sigma_1},\ldots,u_{\sigma_n})$$

has no axiomatic standing. Still worse is true, however. As noted by Sen with reference to the individuals' weights, the *form* of the social welfare function $W(\cdot)$ depends on the chosen profile $(u_1,\ldots,u_n)$; so $(*)$ could only hold by mere chance. Procedure (ii) is not satisfactory either. If it is to have weight as an argument for utilitarianism, the Aggregation Theorem should not make the shape of the social choice rule dependent on changes in the functional representations of utilities that have not been delineated in the axiom system. Within a single-profile approach, there is no way in which (ii) can be fitted to this requirement. In order to go from a weighted sum rule to some symmetric rule, one must independently rescale the individual utility functions. This is a permissible move only if there is an axiom stating that the social preference is invariant with respect to inde-

[9] This claim is probably implicit in the social choice literature of the 70's. However, we do not know any explicit reference. A version of this claim involving specific assumptions can be found in the theory of Arrovian social welfare functions on economic domains: see Example 6 in Le Breton and Weymark (forthcoming).

pendent rescalings. Such an axiom cannot be expressed in the single-profile approach.[10]

In brief, the standard, single-profile approach to the Aggregation Theorem is in trouble when it comes to the issue of symmetry. This is not a problem for the multi-profile approach, since the latter makes it possible to state axiom (A). A word should be said concerning the intermediate case between the single- and multi-profile approaches. It would seem that a suitably enlarged single-profile approach, as exemplified by K. Roberts's work (1980b), could as well accommodate (A). The general idea is to define a social welfare *functional* on a domain D_U which is built up from a particular profile U and is not normally as large as the set of *all* profiles. Starting from $U=(u_1,\dots,u_n)$, the first thing to do is obviously to include in D_U all permuted vectors $(u_{\sigma_1},\dots,u_{\sigma_n})$. More profiles should be added also, for D_U must be rich enough to sustain the same kind of interpolation arguments as were used in the proof of Lemma 1. Whatever the algebraic restrictions finally selected, they will (perhaps) logically weaken the multiprofile assumption $D=\mathscr{L}(\mathscr{M})^n$, but will (most surely) appear less natural. This is the fundamental methodological problem with the enlarged single-profile approach. Nevertheless, if only for the sake of making comparisons easier, it would be interesting to apply this approach to the assumptions of the Aggregation Theorem, as it was applied to more standard assumptions in social choice theory (see d'Aspremont 1985; Sen 1986, for surveys of the main results).

The issue of *interpersonal comparisons of utility* is next on the agenda. The standard formalism for this problem is provided by invariance principles, which can only be stated within a multi-profile or an enlarged single-profile framework. Thus, we could repeat one of the points just made with respect to anonymity: the language in which the initial Aggregation Theorem is phrased is not rich enough to express all of the purported interpretations. This is not to say that there have been no good *informal* discussions of the problem at hand (see below); but they should be easier to understand against the background of Propositions 1 and 2 than of the initial result. As far as interpersonal comparisons are concerned, the salient point is Lemma 4: given any axiom system which implies welfarism, it is equivalent to assume that the social observer satisfies VNM independence or that he compares utilities in the (CU) way. Hence, it would be redundant to combine any of the invariance principles that have been used to formalise utilitarianism - (DC), (CC), or (CU) - with the axioms of either Proposition 1 or 2. Does this mean that the Aggregation Theorem *justifies* either the desirability or the feasibility of utilitarian interpersonal utility comparisons? Prima facie, it does in the following weak, conditional sense: *if* somebody accepts the welfarism axioms and also believes that the social observer does (should) respond to risk in the way recommended by VNM theory, *then* this individual in effect believes that a certain kind of utility comparison - and only that kind - is possible (resp. desirable). Note that it is not so important to distinguish between the "feasibility" and "desirability" of interpersonal comparisons in this context, as the Aggregation Theorem is used only to impose logical coherence on a *system of beliefs:* the latter may just as consistently be concerned with the *is* as with the *ought.*

[10] Formally, such an axiom could be expressed in the multi-profile language, but one might wonder whether it would then be desirable on theoretical grounds. Note that Harsanyi's derivation of a symmetric rule in (1977a, Chap 4.1) involves method (i). Other writers (such as Broome 1991, Chap 10) complement the derivation of the weighted sum formula with some application of (ii).

There are a number of relevant variants of the discursive strategy just sketched. The most defensible ones strike us as being of the negative sort; that is, they trade upon the expression "and only that kind" in the sentence above. For instance, Proposition 1 could be used against anybody who would be prepared to swallow (I) and (SP), make VNM assumptions on the individuals' and the observer's utilities - allegedly because these assumptions reflect individual rationality - and, say, turn Rawlsian, or hostile to any interpersonal comparison whatsoever, when it comes to assessing income distribution. The negative rhetorical use of the Aggregation Theorem amounts to interpreting it as a source of impossibility results rather than a positive foundation for utilitarianism.

Even this unassuming interpretation of the ethical content of the Aggregation Theorem runs into prima facie difficulties. Many of Harsanyi's commentators appear to believe that the possibility of interpersonal comparisons of utility is as much of an assumption of the Theorem as it is one of its conclusions. There is - it is said - an underlying circularity in the reasoning which should prevent one from using it as an argument, even in the negative strategy sketched above. Broome (1991, p 219–220) has usefully discussed this objection. He claims that the possibility of interpersonal comparisons is implicit in the requirement that the social preference be *complete.* The latter is part of the definition of a SWFL in the framework of this paper. Broome's argument may be reexpressed as follows. If the observer can rank any pair x, y in the prospect set, he can indeed rank those pairs x, y which make i better off in x and j better off in y. Once welfarism is taken for granted, the connection between assuming completeness and assuming that any utility comparison whatever is possible becomes obvious. This argument for locating the hidden assumption also provides an elegant solution to the circularity objection. For what is assumed is just the *generic* possibility of making interpersonal comparisons: what is logically derived, and heuristically added, is the *specific*, utilitarian way of making them. The step from the philosophical premises to the conclusion is not as big as it first seemed; but the latter retains some kind of surprise effect compared with the former. This is what makes the Aggregation Theorem an effective argument, although by no means a decisive one.[11]

We agree with Broome that one should distinguish between the possibility of making interpersonal comparisons and their utilitarian nature, and that the contribution of the theorem is to fill the gap between the two notions. But we somehow differ about the suitable location of the hidden assumption. The statement that "interpersonal comparisons of utility are possible" can be understood in more than one way. If it means that the social observer is able to rank *any* pair of utility vectors $U(x)$, $U(y)$, then completeness is indeed crucial, as Broome contends. But the statement can also be understood as saying that the social observer is able to rank *some* pairs $U(x)$, $U(y)$, and that his (generally partial) ranking is internally consistent. For this to be the case the nonemptiness and transitivity of the social preference relation are crucial. Also, the statement under scrutiny has presumably the connotation that interpersonal comparisons are not only possible, but *actually used* in the social assessment of alternatives x, y. This last meaning is implied in the welfarism step on which the proof of the Aggre-

[11] This intepretation seems to run counter to Harsanyi's own view that the "theorem does not depend on the possibility of interpersonal utility comparisons" (1979, p 294). See also the related comment in (1978, p 227).

gation Theorem depends. (As explained above, Lemma 1 is a welfarism lemma in the sense of standard social choice; the corresponding step in the single-profile approach is the statement that Pareto Indifference is equivalent to the factoring out of social utility in terms of individual utilities.) Now, if welfarism turns out to be relevant to the formalisation of "the possibility of interpersonal comparisons", one should take into consideration other components than the relational properties of social preference - in particular the assumption that the observer's datum is a *utility function profile* rather than a preference profile or a profile of classes of utility functions. In sum, there does not appear to be any clear-cut way of locating the assumption that "interpersonal comparisons of utility are possible", although Broome as well as other critiques of Harsanyi's approach were certainly right to claim that it is hidden in the premisses.[12]

We have not yet addressed Sen's reason (1) - see above Sect. 1 - for discarding the initial Aggregation as a *proper* axiomatisation of utilitarianism. This is a very different objection from those just expressed. For Sen is not arguing that the theorem is an incomplete or question-begging argument for utilitarian ethics; rather, that it misses its purpose altogether and contains no argument at all. Part of Sen's claim that there should be an "independent formulation of individual utilities" points towards a semantic elucidation of the utility concept suitable for utilitarianism. This line of inquiry is present in Weymark's (1991) review of the Sen-Harsanyi debate and pursued thoroughly in Broome (1991). It goes far beyond the scope of this paper, which tries to remain agnostic on the ultimate interpretation of "utility". There is, however, a more circumscribed analysis of reason (1) which should be mentioned here. *The utilitarian-looking conclusion of the initial Aggregation Theorem depends on assuming particular types of utility functions to represent the observer's and the individuals' preferences. These assumptions might prove difficult to justify.* Weymark (1991, p 297–312) particularly focuses on these kinds of criticisms.[13] At stake here are the following three problems: (i) Why should the social theorist select a MP representation of the observer's VNM preferences and assume that the observer only relies on MP representations of the individuals' VNM preferences? (ii) Is the cardinalisation of individual utility that the VNM preference axioms imply at all relevant to the derivation of utilitarianism? That is to say, do the VNM axioms deliver a procedure to measure the intensity of individual preferences, as they should if they bear on the derivation of utilitarianism? (iii) Should the social choice theorist have in the first place assumed the VNM preference axioms to hold of the observer and the individuals?

Question (iii) is of course the reason for the most strongly voiced disagreements among commentators. Harsanyi notwithstanding, today's decision and social choice theorists do not accept anymore the equation of rationality with the VNM independence axiom as being unproblematic. Allais as well as many others after him have taken von Neumann and Morgenstern to task for having neglected the

---

[12] To clarify the role of the completeness property of social preference it would be interesting to try a variant of Proposition 1 which dispenses with this assumption. As mentioned in Sect. 2.3, our welfarism lemma does not depend on it; nor does Lemma 4. This point would give support to the following variant of the interpretation provided earlier in the text: "if somebody accepts the welfarism axioms and also believes that the social observer does respond to risk in the way recommended by VNM theory, then this individual in effect believes that only a certain kind of utility comparisons [= CU] take place *whenever utility comparisons are at all possible*".

[13] For an early discussion of related issues see Blackorby et al. (1980).

impact of the dispersion of utility values across outcomes.[14] There is a ruling debate on the alleged connection between either VNM independence or Savage's sure-thing principle and dynamic consistency.[15] It would be a gross mistake to believe that the critiques of expected utility have confined their case to empirical issues. Since they also take sides on *normative* issues, a complete assessment of the Aggregation Theorem cannot dispense with some view of the current controversies. However, this paper does not aim at such a complete assessment. Sen's last-mentioned objection points towards questions (i) and (ii) rather than (iii). We shall examine them from the angle of the VNM theory, relying on Weymark's important logical point that *even* supporters of this theory should address them. For one, the VNM existence and uniqueness theorem does not force the use of MP functions to represent the agent's binary choices. For another, the relative uniqueness of MP representations does not in itself imply that they provide a meaningful way of measuring preference differences.

In this writer's opinion consistent supporters of the VNM preference axioms should be able to answer (i) more easily than (ii). In connection with (i) note carefully that the multi-profile reconstruction of the Aggregation Theorem does not use any MP representation *on the observer's part.* Rather, it is phrased in terms of his VNM preferences over lotteries. This would suggest to explore a corresponding single-profile variant of the theorem; it might be the case that the relational arguments underlying Lemma 2 and the end of the proof of Proposition 1 just carry through. It is a fact that both the single- and multi-profile versions of the Aggregation Theorem crucially depend on considering only MP representations *on the individuals' part.* But there are arguments rooted in the VNM theory at large to defend this restriction. Assuming a set of monetary outcomes and that the VNM axioms hold of the particular individual, there are simple - and well-known - procedures to elicit a utility function on the set of sure outcomes, using his observed choices among lotteries as data. Having elicited one such function for each individual, the ideal observer can of course extend it into a MP utility function on the set of lotteries. Now, one reason why he may restrict himself to affine positive transforms of that particular representation is that any other ordinal transforms would change the individual's Arrow-Pratt coefficient of risk aversion. That is, the observer would loose any information on risk attitudes; he would just preserve information on the binary choices made by the individual. Presumably, the social observer is interested in recording not only the latter but also the former. Harsanyi's motivation at various stages of his ethical theory has been to connect individual attitudes towards risk with social attitudes towards inequality.[16]

Even writers leaning towards the VNM theory might be in serious trouble when facing problem (ii). If the additive formula derived by the Aggregation

---

[14] See Allais's classic 1953 paper and a response to it from the viewpoint of VNM theory in Broome, 1991, p 110. Note that this side of the discussion interacts with the debate over distributional consequences of utilitarianism.

[15] See Hammond's (1988) reconstruction of independence and the sure-thing principle using "consequentialism". McClennen (1990) provides a thorough analysis of Hammonds's argument, as well as an alternative account of dynamic rationality leading to a skeptical view of independence.

[16] Admittedly, the argument of this paragraph is a heuristic one. To formalise it, one migh have to redefine SWFL's so as to make them explicitly dependent on the individuals' Arrow-Pratt coefficients of risk aversion.

Theorem is to count as truly utilitarian, it must be amenable to the standard philosophical interpretation of utilitarianism, to the effect that the observer's preference between x and y depends on comparing the intensity of preference of x over y in one subgroup of individuals with the intensity of the opposite preference in the complementary subgroup. This unproblematic point appears to clash with the widespread view that one should distinguish between *two senses of cardinality* when dealing with the VNM theory. MP representations of the individuals' VNM preferences are cardinal in the sense of being unique up to positive affine transforms but not necessarily cardinal in the more relevant sense of measuring the intensity of preferences. At an early stage Luce and Raiffa (1957, p 32) warned users of the VNM theory against the "fallacy" of believing that the former sense *logically implied* the latter. This warning has become part of the standard teaching.[17] Not only do contemporary writers in decision theory take Luce and Raiffa's warning seriously, but a majority of them have concluded that MP indexes actually do *not* represent the intensity of preferences. Harsanyi here dissents from the mainstream: "even though a person's VNM utility function is always estimated in terms of his behaviour under risk and uncertainty, the real purpose of this estimation procedure is to obtain cardinal-utility measures for the relative importance he assigns to various economic (and noneconomic) alternatives" (1977b, in 1982, p 53). To arbitrate this conflict between Harsanyi and the current VNM doctrine would again lead us beyond the scope of this essay. We shall just mention that only a *normative* claim could fill the gap between the mathematical sense of cardinality and the conceptually relevant one. For there is no logical bridge, as Luce and Raiffa explained and Weymark reminded. There is no factual bridge either, because evidence is by and large unfavourable to the claim that MP functions adequately measure preference differences, as independently revealed to the experimenter.[18] Question (ii) is perhaps the most vexing one among those left unsolved by the present reconstruction of the Aggregation Theorem.

References

d'Aspremont C (1985) Axioms for social welfare orderings. In: Hurwicz L, Schmeidler D, Sonnenschein H (eds) Social Goals and Social Organization. Cambridge: Cambridge University Press

d'Aspremont C, Gevers L (1977) Equity and the informational basis of collective choice. Rev Econ Stud 44: 199–209

Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. Econometrica 21: 503–546

Blackorby C, Donaldson D (1982) Ratio-scale and translation-scale full interpersonal comparability without domain restrictions. admissible social-evaluation functions. Int Econ Rev 23: 249–268

Blackorby C, Donaldson D, Weymark JA (1980) On John Harsanyi's defences of utilitarianism. CORE Discussion Paper 8013, Université Catholique de Louvain

Blackorby C, Donaldson D, Weymark JA (1984) Social choice with interpersonal utility comparisons: a diagrammatic introduction. Int Econ Rev 25: 327–356

Blackwell D, Girschick MA (1954) Theory of games and statistical decisions. New York: Wiley

[17] E.g. Fishburn (1970, p 80–82). See Fishburn (1989) and Bouyssou and Vansnick (1990) for surveys of the measurability issue in expected-utility theory.

[18] We rely here on the evidence surveyed in Bouyssou and Vansnick (1990).

Bouyssou D, Vansnick JC (1990) "Utilité cardinale" dans le certain et choix dans le risque. Rev économique 6: 979–1000
Broome J (1987) Utilitarianism and expected utility. J Philos 84: 405–422
Broome J (1990) Bolker-Jeffrey expected utility theory and axiomatic utilitarianism. Rev Econ Stud 57: 477–502
Broome J (1991) Weighing Goods. Oxford: Basil Blackwell
Coulhon T, Mongin P (1989) Social choice theory in the case of von Neumann-Morgenstern utilities. Soc Choice Welfare 6: 175–187
De Meyer B, Mongin P (1994) A note on affine aggregation. CORE Discussion Paper 9414, Université Catholique de Louvain
Deschamps R, Gevers L (1978) Leximin and utilitarian rules: a joint characterization. J Econ Theory 17: 143–163
Deschamps R, Gevers L (1979) Separability, risk-bearing and social welfare judgments. In: Laffont JJ (ed) Aggregation and Revelation of Preferences. Amsterdam: Nort-Holland
Diamond PA (1967) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: comment. J Pol Econ 75: 765–766
Fishburn PC (1970) Utility theory for decision-making. New York: Wiley
Fishburn PC (1982) The foundations of expected utility. Dordrecht: Reidel
Fishburn PC (1984) On Harsanyi's utilitarian cardinal welfare theorem. Theory Decis 17: 21–28
Fishburn PC (1989) Retrospective on the utility theory of von Neumann and Morgenstern. J Risk Uncert 2: 127–158
Gevers L (1979) On interpersonal comparability and social welfare orderings. Econometrica 47: 75–89
Halmos PR (1950) Measure Theory. Princeton: van Nostrand
Hammond PJ (1982) Utilitarianism, uncertainty and information. In: Sen A, Williams B (eds) Utilitarianism and Beyond. Cambridge: Cambridge University Press
Hammond PJ (1987) On reconciling Arrow's theory of social choice with Harsanyi's fundamental utilitarianism. In: Feiwel GR (ed) Arrow and the Foundations of the Theory of Economic Policy. London: MacMillan
Hammond PJ (1988) Consequentialist foundations for expected utility. Theory Decis 25: 25–78
Hammond PJ (1991) Interpersonal comparisons of utility: why and how they are and should be made. In: Elster J, Roemer JE (eds) Interpersonal Comparisons of Well-Being. Cambridge: Cambridge University Press
Harsanyi JC (1953) Cardinal utility in welfare economics and in the theory of risk-taking. J Polit Econ 61: 434–435
Harsanyi JC (1955) Cardinal welfare, individualistic ethics and interpersonal comparisons of utility. J Polit Econ 63: 309–321
Harsanyi JC (1975) Nonlinear social welfare functions: do welfare economists have a special exemption from Bayesian rationality? Theory Decis 6: 311–332
Harsanyi JC (1977a) Rational behavior and bargaining equilibrium in games and social situations. Cambridge: Cambridge University Press
Harsanyi JC (1977b) Morality and the theory of rational behavior. Social Res 44: 623–656; Reprinted (1982). In: Sen A, Williams B (eds) Utilitarianism and Beyond. Cambridge: Cambridge University Press
Harsanyi JC (1978) Bayesian decision theory and utilitarian ethics. Am Econ Rev (Papers and Proceedings) 68: 223–228
Harsanyi JC (1979) Bayesian decision theory, rule utilitarianism, and Arrow's impossibility theorem. Theory Decis 11: 289–317
Harsanyi JC (1987) Von Neumann-Morgenstern utilities, risk taking, and welfare. In: Feiwel GR (ed) Arrow and the ascent of Modern Economic Theory. London: Mac Millan
Kelly PJ (1979) Geometry and Convexity. New York: Wiley
Le Breton M, Weymark J (forthcoming) An introduction to Arrovian social welfare functions on economic and political domains. In: Schofield N (ed) Social Choice and Political Economy. Dordrecht: Kluwer
Luce RD, Raiffa H (1957) Games and decisions. New York: Wiley
Maskin E (1978) A theorem on utilitarianism. Rev Econ Stud 45: 93–96

McClennen EF (1990) Rationality and dynamic choice. Cambridge: Cambridge University Press

Mongin P (1990) Utilitarianism, invariance principles, and the von Neumann-Morgenstern hypothesis. Document 90A13 GREQE, Ecole des Hautes Etudes en Sciences Sociales, Marseille

Mongin P (1993) Consistent Bayesian aggregation. CORE Discussion Paper 9319, Université Catholique de Louvain, forthcoming in J Econ Theory

Myerson RB (1981) Utilitarianism, egalitarianism, and the timing effect in social choice problems. Econometrica 49: 883-897

Roberts KWS (1980a) Interpersonal comparability and social choice theory. Rev Econ Stud 47: 421-439

Roberts KWS (1980b) Social choice theory: the single and multi-profile approaches. Rev Econ Stud 47: 441-450

Sen AK (1970) Collective choice and social welfare: San Francisco: Holden Day

Sen AK (1974) Informational bases of alternative welfare approaches: aggregation and income distribution. J Publ Econ 3: 387-403

Sen AK (1986) Social choice theory. In: Arrow KJ, Intriligator MD (eds) Handbook of Mathematical Economics III p 1073-1181 Amsterdam: North-Holland

Weymark JA (1991) A reconsideration of the Harsanyi-Sen debate on utilitarianism. In: Elster J, Roemer J (eds) Interpersonal Comparisons of Well-Being. Cambridge: Cambridge University Press

Weymark JA (1993) Harsanyi's social aggregation theorem and the weak Pareto principle. Soc Choice Welfare 10: 209-221

Weymark JA (1994) Harsanyi's social aggregation theorem with alternative Pareto principles. In: Eichhorn W (ed) Models and Measurement of Welfare and Inequality. Berlin, Heidelberg, New York: Springer