

The Informational Basis of the Theory of Fair Allocation*

Marc Fleurbaey[†] Kotaro Suzumura[‡] Koichi Tadenuma[§]

This version: November 2002
(First version: July 2001)

Abstract

The theory of fair allocation is often favourably contrasted with the social choice theory in the search for escape routes from Arrow's impossibility theorem. Its success is commonly attributed to the fact that it is modest in its goal vis-à-vis social choice theory, since it does not aspire for a full-fledged ordering of options, and settles with a subset of "fair" options. We show that its success may rather be attributable to a broadened informational basis thereof. To substantiate this claim, we compare the informational basis of the theory of fair allocation with the informational requirements of social choice theory.

JEL Classification Numbers: D63, D71.

*This paper is derived from an earlier draft of our paper entitled "Informational requirements for social choice in economic environments". The authors thank A. Trannoy for comments, and participants at seminars in Cergy and Rochester, and the 5th International Meeting of the Society for Social Choice and Welfare in Alicante. Financial support from the Ministry of Education of Japan through Grant-in-Aid No. 10045010 ("Economic Institutions and Social Norms: Evolution and Transformation") is gratefully acknowledged.

[†]CATT, THEMA, IDEP, University of Pau, Av. du Doyen Poplawski, 64000 PAU (France). Email: Marc.Fleurbaey@univ-pau.fr

[‡]Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo 186-8603 (Japan). Email: suzumura@ier.hit-u.ac.jp

[§]Faculty of Economics, Hitotsubashi University, Kunitachi, Tokyo 186-8601 (Japan). Email: tadenuma@econ.hit-u.ac.jp

1 Introduction

The theory of fair allocation studies allocation rules which select, for every economy in a given class, a subset of feasible allocations on the basis of efficiency and fairness properties. It was initiated by Foley (1967), Kolm (1972) and Varian (1974) among others, who focussed on the concept of no-envy. Since then it has been extended to cover many other notions of fairness and a great variety of economic contexts (production, public goods, etc.) by many authors.¹ This theory contains some negative results, because it is usually impossible to find solutions which satisfy all conceivable requirements of efficiency and equity simultaneously, but its hallmark is a richness of positive results. By now, not only are there many interesting allocation rules uncovered in the literature, but also they are fully characterized as the only rules satisfying some sets of reasonable axioms.

Compared to the theory of social choice, this makes a great contrast. In social choice theory, Arrow's impossibility theorem has been shown to remain valid in most economic or abstract contexts. This theorem, like all the theory of social choice, is about social preferences which rank all options in a given set on the basis of individual preferences over these options. The theorem states that there is no way to construct social preferences as a function of individual preferences if this function is required to satisfy basic principles of unanimity (Weak Pareto: if everybody prefers x to y , so does society), impartiality (Non-Dictatorship: no individual can always impose his strict preferences) and informational parsimony (Independence of Irrelevant Alternatives: social preferences over any subset of alternatives depend only on individual preferences over this subset).

Impossibilities in social choice theory, possibilities in fair allocation theory. This contrast requires an explanation. Most authors have stressed two differences between the two theories. The one most often mentioned is about preferences versus selection. Varian (1974) argues as follows:

‘Social [choice] theory asks for too much out of the process in that it asks for an entire *ordering* of the various social states (allocations in this case). The original question asked only for a “good” allocation; there was no requirement to rank all allocations. The fairness criterion in fact limits itself to answering the original question. It is limited in that it gives no indication of the

¹For a survey, see Moulin and Thomson (1997).

merits of two nonfair allocations, but by restricting itself in this way it allows for a reasonable solution to the original problem.’ (p. 65)

Similarly, Kolm (1996) is ironic about social preferences:

‘The requirement of a social ordering is indeed problematic at first sight: Why would we want to know the 193th best alternative? Only the first best is required for the choice.’ (p. 439)

In his famous survey on social choice theory, Sen (1986) also emphasizes this contrast:

‘The specified subset is seen as good, but there is no claim that they represent the “best” alternatives, all *equally* choosable. There is no attempt to give an answer to the overall problem of social choice, and the exercise is quite different from the specification of a social preference over X .’ (p. 1106)

And most recently, Moulin and Thomson (1997) have compared the two theories in these terms:

‘In social choice theory, the focus is commonly on obtaining a complete ranking of the set of feasible alternatives as a function of the profile of individual preferences. (...) Consider now the axiomatic investigations of resource allocation. As their counterparts in the theory of cooperative games, their focus is on the search for allocation rules, no attempt being made at obtaining a complete ranking of the entire feasible set.’ (p. 104)

The second difference noticed by these authors is that economic models enable the analyst to take account of the structure of allocations. Varian mentions only the fact that the theory of fairness can focus on self-centered preferences (individuals being interested only in their own consumption), while Sen has written about the fairness literature:

‘First, it has shown the relevance of informational parameters that the traditional social choice approaches have tended to ignore in the single-minded concern with individual orderings of

complete social states. Comparisons of different persons' positions within a state have been brought into the calculation, enlarging the informational basis of social judgments. Second, in raising rather concrete questions regarding states of affairs, the fairness literature has pushed social choice theory in the direction of more structure.' (p. 1111)

Similarly, Moulin and Thomson have argued that

'the models of resource allocation take full account of the microeconomic structure of the problems to be solved. (...) This descriptive richness permits a great deal of flexibility at two levels. First, properties of allocation rules can be formulated directly in terms of the physical attributes of the economy (...). Second, the rich mathematical structure of microeconomic models gives rise to a host of variations on each general principle.' (p. 105)

However, this second difference is about additional requirements formulated in a richer framework, and can hardly explain the relative success of the theory of fairness. This was noted by Moulin and Thomson, who have concluded:

'Note that social choice theory itself has recently developed in a similar direction, widening its framework by incorporating information about economic environments (...). But as its objective has remained to obtain complete rankings of sets of feasible alternatives, its conclusions have so far remained largely negative.' (ibid.)

Actually, Arrow's initial presentation of his theorem (Arrow 1950, 1951) was already formulated in an economic setting, with self-centered preferences. Therefore one can safely conclude that the common explanation for the possibility results in the theory of fairness is that it does not seek a full-fledged ordering.

The problem we address in this paper is that this common explanation is technically incomplete and therefore unconvincing. An allocation rule, in effect, splits the set of allocations in two parts, the good and the bad. Even though the ambition is not to give an ordering of allocations, this twofold partition is, *formally*, an ordering. A coarse ordering is still an

ordering. It is a partial ordering if the good allocations are deemed non-comparable, and similarly for the bad ones, as suggested above by Varian and Sen. But one may also consider that it gives a complete ordering simply by declaring all good allocations, and respectively all bad allocations, to be socially equivalent. We will later refer to these two interpretations as the “partial ordering” version and the “complete ordering” version, respectively.

If all the allocation rules in the theory of fairness yield orderings, it remains totally unexplained why and how the theory of fair allocation avoids Arrow’s impossibility. There may be some truth in the idea that a coarse ordering is more easily constructed than a fine-grained one, but one cannot say that a coarse ordering is not an ordering and is outside the reach of Arrow’s theorem.

As a consequence, the only convincing explanation for the possibility results in fair allocation theory must be based on the violation of some conditions of Arrow’s theorem. The first candidate, among Arrow’s conditions, is completeness of the ordering. This is in line with the “partial ordering” interpretation of allocation rules. The problem is that under the “complete ordering” interpretation, allocation rules do provide complete orderings. Therefore this cannot be the desired explanation.² If one can derive complete orderings from the theory of fairness, the explanation must be that this theory violates one of the three core axioms of Arrow’s theorem: Weak Pareto, Non-Dictatorship, or Independence of Irrelevant Alternatives.

These three axioms will be formally defined in the next sections, after the model and notations are introduced. *The main conclusion that will be derived in this paper is that the theory of fairness succeeds in obtaining possibility results mainly because it abandons the axiom of Independence of Irrelevant Alternatives.* It relies on more information about individual preferences (at “irrelevant” alternatives) than allowed by this axiom. This idea was already informally put forth by Fleurbaey and Maniquet (1996), who claimed moreover that the kind of additional information used by the theory of fairness would be enough to get round Arrow’s impossibility in the theory of social choice as well.

Here we will show, however, that relaxing Weak Pareto is also part of the

²Moreover, Weymark (1984) has studied the application of Arrow’s axioms to partial orderings, and obtained oligarchy results. More interestingly, by adding anonymity to the axioms, he characterized the Pareto partial ordering. Although his results are obtained in an abstract framework with unrestricted preferences, they strongly suggest that little can be gained by abandoning completeness.

picture, and this is an inevitable consequence of the fact that the ordering is coarse. Our analysis will therefore show that there is some truth in the usual explanation. More precisely, coarse orderings naturally suggest weakenings of the Weak Pareto principle which make it less necessary to depart from Independence of Irrelevant Alternatives. With the full Weak Pareto principle for social orderings, avoiding Arrow’s impossibility would require a more drastic weakening of the independence condition. Therefore our second important conclusion will be that *the theory of social choice needs more information than the theory of fair allocation*, although both need to relax the independence condition. This part of our analysis will rely heavily on a companion paper which studies how much weakening of the independence condition is required in the theory of social choice (Fleurbaey, Suzumura and Tadenuma 2001).

The paper has the following structure. The next section introduces the model and the main notions. Allocation rules are confronted to Arrow’s axioms in Section 3. In Section 4 we then discuss some weak variants of Independence of Irrelevant Alternatives. Section 5 presents the main results, and Section 6 compares the informational basis of the theory of fairness to that of social choice theory. In Section 7 we propose a unified theory of social choice and fairness, and discuss the relative merits of alternative “grand unification” theories. Section 8 concludes.

2 Model and definitions

2.1 The model

The *population* is fixed. Let $N = \{1, \dots, n\}$ be the set of *agents* where $2 \leq n < \infty$. There are ℓ *goods* indexed by $k = 1, \dots, \ell$ where $2 \leq \ell < \infty$. Agent i ’s *consumption bundle* is a vector $x_i = (x_{i1}, \dots, x_{i\ell}) \in \mathbb{R}_+^\ell$. An *allocation* is denoted $x = (x_1, \dots, x_n) \in \mathbb{R}_+^{n\ell}$.

A *preordering* is a reflexive and transitive binary relation. Agent i ’s *preferences* are described by a complete preordering R_i (strict preference P_i , indifference I_i) on \mathbb{R}_+^ℓ . A *profile of preferences* is denoted $\mathbf{R} = (R_1, \dots, R_n)$. Let \mathcal{R} be the set of continuous, convex, and strictly monotonic preferences over \mathbb{R}_+^ℓ .

Let π be a bijection on N . For each $x \in \mathbb{R}_+^{n\ell}$, define $\pi(x) = (x'_1, \dots, x'_n) \in \mathbb{R}_+^{n\ell}$ by $x'_i = x_{\pi(i)}$ for all $i \in N$, and for each $\mathbf{R} \in \mathcal{R}^n$, define $\pi(\mathbf{R}) =$

$(R'_1, \dots, R'_n) \in \mathcal{R}^n$ by $R'_i = R_{\pi(i)}$ for all $i \in N$. Let Π be the set of all bijections on N .

There is no production in our model, and the amount of *total resources* is fixed and represented by the vector $\omega \in \mathbb{R}_{++}^\ell$. An allocation $x \in \mathbb{R}_+^{n\ell}$ is *feasible* if $\sum_{i \in N} x_i \leq \omega$.³ Let F be the set of all feasible allocations.

For each $\mathbf{R} \in \mathcal{R}^n$, let $E(\mathbf{R})$ denote the set of *Pareto-efficient allocations*. Because of strict monotonicity of preferences, there is no need to distinguish Pareto-efficiency in the strong sense and in the weak sense.

A *social ordering function* (SOF) is a function \bar{R} defined on \mathcal{R}^n , such that for all $\mathbf{R} \in \mathcal{R}^n$, $\bar{R}(\mathbf{R})$ is a complete preordering on the set of allocations F . Let $\bar{P}(\mathbf{R})$ (resp. $\bar{I}(\mathbf{R})$) denote the strict preference (resp. indifference) relation derived from $\bar{R}(\mathbf{R})$.

An *allocation rule* (AR) is a set-valued mapping S defined on \mathcal{R}^n , such that⁴ for all $\mathbf{R} \in \mathcal{R}^n$, $S(\mathbf{R})$ is a non-empty subset of F . An AR S is *essentially single-valued* if all selected allocations are Pareto-indifferent:

$$\forall x, y \in S(\mathbf{R}), \forall i \in N, x_i I_i y_i.$$

To each AR S can be associated the (two-tier) SOF \bar{R}_S defined as follows: for all $\mathbf{R} \in \mathcal{R}^n$, and all $x, y \in F$,

$$x \bar{R}_S(\mathbf{R}) y \Leftrightarrow x \in S(\mathbf{R}) \text{ or } y \notin S(\mathbf{R}).$$

One then has: for all $\mathbf{R} \in \mathcal{R}^n$, and all $x, y \in F$,

$$x \bar{P}_S(\mathbf{R}) y \Leftrightarrow \text{not}[y \bar{R}_S(\mathbf{R}) x] \Leftrightarrow x \in S(\mathbf{R}) \text{ and } y \notin S(\mathbf{R}).$$

When a SOF has only two indifference classes, and is therefore associated to an AR, we will call it an ARSOF. We will say that an ARSOF is *essentially single-valued* if its associated AR is essentially single-valued.

Conversely, from any SOF \bar{R} , we can derive the AR $S_{\bar{R}}$ by selecting the subset of first best allocations: for all $\mathbf{R} \in \mathcal{R}^n$,

$$S_{\bar{R}}(\mathbf{R}) = \{x \in F \mid \forall y \in F, x \bar{R}(\mathbf{R}) y\}.$$

By the definitions, we have, for any AR S :

$$S_{\bar{R}_S} = S.$$

³Vector inequalities are denoted as usual: \geq , $>$, and \gg .

⁴An alternative definition of SOFs and ARs makes them a function of ω as well as \mathbf{R} . This is useful when changes in ω are studied, but here we focus only on the information about preferences, and since ω is kept fixed throughout the paper, we omit this argument.

2.2 Arrow's axioms

We are now ready to give precise definitions of Arrow's three conditions.

Weak Pareto: $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N$, $x_i P_i y_i$, then $x \bar{P}(\mathbf{R}) y$.

Independence of Irrelevant Alternatives (IIA): $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N$,

$$\begin{aligned} x_i R_i y_i &\Leftrightarrow x_i R'_i y_i \\ y_i R_i x_i &\Leftrightarrow y_i R'_i x_i, \end{aligned}$$

then $x \bar{R}(\mathbf{R}) y \Leftrightarrow x \bar{R}(\mathbf{R}') y$.

In economic domains, it is common to refine the definition of non-dictatorship so as to allow for slight strengthenings of the usual axiom. Let $X \subset F$ be any subset of allocations.

Non-Dictatorship (over X): There does not exist $i_0 \in N$ such that:

$$\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in X, x_{i_0} P_{i_0} y_{i_0} \Rightarrow x \bar{P}(\mathbf{R}) y.$$

In addition to these axioms, it will be useful to refer to a full anonymity condition, which is stronger than Non-Dictatorship but quite appealing on grounds of impartiality:

Anonymity: $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F, \forall \pi \in \Pi$,

$$x \bar{R}(\mathbf{R}) y \Leftrightarrow \pi(x) \bar{R}(\pi(\mathbf{R})) \pi(y).$$

3 Allocation rules and Arrow's axioms

In this section, we examine how ARSOFs, as a particular kind of SOFs, fare with respect to satisfying Arrow's axioms.

Let us start with the last one. The Anonymity axiom raises no difficulty to ARSOFs, and it is worth noticing that it then actually boils down to the following simple condition:

Anonymity for $S_{\bar{R}}$: $\forall \pi \in \Pi, \forall \mathbf{R} \in \mathcal{R}^n, \forall x \in S_{\bar{R}}(\mathbf{R}), \pi(x) \in S_{\bar{R}}(\pi(\mathbf{R}))$.

It is interesting that, moreover, the Non-Dictatorship axiom will (for sufficiently large X) always be trivially satisfied by ARSOFs, since it is impossible for any agent to impose his fine-grained preferences to a coarse social ordering.

The Weak Pareto principle, by contrast, cannot be satisfied by ARSOFs. Since there usually are many occurrences of Pareto-domination among Pareto inefficient allocations, this axiom requires a too fine-grained ranking of allocations. Hence, ARSOFs usually satisfy only weakenings of the Weak Pareto principle. Notice first that the original Weak Pareto principle for ARSOFs can be written as:

Weak Pareto for ARSOF: $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N, x_i P_i y_i$, then $x \in S_{\bar{R}}(\mathbf{R})$ and $y \notin S_{\bar{R}}(\mathbf{R})$.⁵

A priori, one may consider three kinds of weakenings of this original axiom, depending on whether the conclusion of the weakened axiom retains:

- (α) $y \notin S_{\bar{R}}(\mathbf{R})$
- (β) $x \in S_{\bar{R}}(\mathbf{R})$
- (γ) $x \in S_{\bar{R}}(\mathbf{R})$ or $y \notin S_{\bar{R}}(\mathbf{R})$.

Note that

- (i) (α) \Rightarrow (γ), and (β) \Rightarrow (γ), and
- (ii) (γ) $\Leftrightarrow x \bar{R}(\mathbf{R}) y$.

This yields three different axioms, which we now examine in turn.

Weak Weak Pareto α : $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N, x_i P_i y_i$, then $y \notin S_{\bar{R}}(\mathbf{R})$.

This axiom is the most relevant weakening of Weak Pareto for our purposes, because it is equivalent to one of the fundamental axioms in the theory of fair allocation, namely:

Pareto-Efficiency: $\forall \mathbf{R} \in \mathcal{R}^n, S_{\bar{R}}(\mathbf{R}) \subseteq E(\mathbf{R})$.

The second weakening of Weak Pareto is the following:

Weak Weak Pareto β : $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N, x_i P_i y_i$, then $x \in S_{\bar{R}}(\mathbf{R})$.

This is equivalent to: $\forall \mathbf{R} \in \mathcal{R}^n$,

$$S_{\bar{R}}(\mathbf{R}) \supseteq \{x \in F \mid \exists y \in F, \forall i \in N, x_i P_i y_i\}.$$

That is, $S_{\bar{R}}(\mathbf{R})$ must contain all feasible allocations that Pareto dominate *some* feasible allocations. This is not reasonable. The third weakening,

⁵Recall that

$$x \bar{P}(\mathbf{R}) y \Leftrightarrow x \in S_{\bar{R}}(\mathbf{R}) \text{ and } y \notin S_{\bar{R}}(\mathbf{R}).$$

however, is logically weaker than Weak Weak Pareto α and deserves some attention:

Weak Weak Pareto γ : $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N$, $x_i P_i y_i$, then $x \in S_{\bar{R}}(\mathbf{R})$ or $y \notin S_{\bar{R}}(\mathbf{R})$.

This is equivalent to the following condition: $\forall \mathbf{R} \in \mathcal{R}^n, \forall x, y \in F$, if $\forall i \in N$, $x_i P_i y_i$, then $y \in S_{\bar{R}}(\mathbf{R}) \Rightarrow x \in S_{\bar{R}}(\mathbf{R})$. If an ARSOF \bar{R} satisfies this condition, then for each $\mathbf{R} \in \mathcal{R}^n$, there exists $T \subseteq F$ such that

$$S_{\bar{R}}(\mathbf{R}) \supseteq \bigcup_{y \in T} [\{y\} \cup \{x \in F \mid \forall i \in N, x_i P_i y_i\}].$$

If $T \subseteq E(\mathbf{R})$, then $S_{\bar{R}}(\mathbf{R}) \subseteq E(\mathbf{R})$, which is the same implication as Weak Weak Pareto α . If $T \supseteq F$, then $S_{\bar{R}}(\mathbf{R}) \supseteq \{x \in F \mid \exists y \in F, \forall i \in N, x_i P_i y_i\}$, the same implication as Weak Weak Pareto β .

Weak Weak Pareto α and Weak Weak Pareto γ have been introduced in Suzumura (1980), under the denominations of Exclusion Pareto and Inclusion Pareto, respectively. There are interesting ARSOFs satisfying Weak Weak Pareto γ but not Weak Weak Pareto α . For instance, define

$$S_{\bar{R}}(\mathbf{R}) \equiv \{x \in F \mid \forall i \in N, x_i R_i \frac{\omega}{n}\},$$

i.e., $S_{\bar{R}}(\mathbf{R})$ is the set of individually rational allocations from the equal division of resources. This rule satisfies Weak Weak Pareto γ , with $T = \{(\frac{\omega}{n}, \dots, \frac{\omega}{n})\}$.

Let us finally consider IIA. For an ARSOF, this axiom is quite demanding, since it requires that if an allocation is selected while another is not, this does not change when individual preferences relative to these two allocations remain the same, independently of preferences over other allocations. It may be useful to write down this condition explicitly in order to make this point clear:

IIA for ARSOFs: $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n, \forall x, y \in F$,
if $\forall i \in N$,

$$\begin{aligned} x_i R_i y_i &\Leftrightarrow x_i R'_i y_i \\ y_i R_i x_i &\Leftrightarrow y_i R'_i x_i, \end{aligned}$$

then $[x \in S_{\bar{R}}(\mathbf{R}) \text{ and } y \notin S_{\bar{R}}(\mathbf{R})] \Leftrightarrow [x \in S_{\bar{R}}(\mathbf{R}') \text{ and } y \notin S_{\bar{R}}(\mathbf{R}')]$.

For ARSOFs, the above axiom is rigorously equivalent to the original IIA axiom. This may be understood by checking that IIA can equivalently be written with the conclusion

$$x \bar{P}(\mathbf{R}) y \Leftrightarrow x \bar{P}(\mathbf{R}') y$$

instead of

$$x \bar{R}(\mathbf{R}) y \Leftrightarrow x \bar{R}(\mathbf{R}') y$$

and recalling that

$$x \bar{P}(\mathbf{R}) y \Leftrightarrow x \in S_{\bar{R}}(\mathbf{R}) \text{ and } y \notin S_{\bar{R}}(\mathbf{R}).$$

4 IIA and Weak Weak Pareto

The fact that IIA for ARSOFs is very strong is captured in the following result, which says that it implies that social preferences are totally independent of individual preferences.

Proposition 1 *An ARSOF \bar{R} satisfies IIA if and only if $S_{\bar{R}}$ is a constant function.*

Proof. It is obvious that a constant ARSOF satisfies IIA. For the converse, choose $i_0 \in N$ and define $x^0 \in F$ by $x_{i_0}^0 = \omega$ (and $x_i^0 = 0$ for all $i \neq i_0$). If for all $\mathbf{R} \in \mathcal{R}^n$ one has $S(\mathbf{R}) = F$, then S is a constant function. Suppose then that this is not the case, and let $\mathbf{R} \in \mathcal{R}^n$ be such that $S(\mathbf{R}) \neq F$.

First case: $x^0 \in S(\mathbf{R})$. Take any $y \notin S(\mathbf{R})$. By monotonicity of preferences, for all $\mathbf{R}' \in \mathcal{R}^n$,

$$\forall i \in N, x_i^0 R_i y_i \Leftrightarrow x_i^0 R'_i y_i \text{ and } y_i R_i x_i^0 \Leftrightarrow y_i R'_i x_i^0.$$

Therefore $x^0 \in S(\mathbf{R}')$ and $y \notin S(\mathbf{R}')$. The latter implies $F \setminus S(\mathbf{R}) \subset F \setminus S(\mathbf{R}')$. Since $x^0 \in S(\mathbf{R}')$, one can show by a symmetrical argument that $F \setminus S(\mathbf{R}') \subset F \setminus S(\mathbf{R})$ implying $S(\mathbf{R}') = S(\mathbf{R})$.

Second case: $x^0 \notin S(\mathbf{R})$. Take any $x \in S(\mathbf{R})$. By monotonicity of preferences, for all $\mathbf{R}' \in \mathcal{R}^n$,

$$\forall i \in N, x_i^0 R_i x_i \Leftrightarrow x_i^0 R'_i x_i \text{ and } x_i R_i x_i^0 \Leftrightarrow x_i R'_i x_i^0.$$

Therefore $x^0 \notin S(\mathbf{R}')$ and $x \in S(\mathbf{R}')$. Hence, $S(\mathbf{R}) \subset S(\mathbf{R}')$. Similarly, by a symmetrical argument based on $x^0 \notin S(\mathbf{R}')$, one can show that $S(\mathbf{R}') \subset S(\mathbf{R})$. ■

Contrary to what one might expect, this does not exactly entail an Arrowian impossibility. In fact, there are ARSOFs satisfying IIA and Weak Weak Pareto principles. Let us first examine the implication of IIA together with the weakest of our Weak Pareto principles, namely Weak Weak Pareto γ . Let F^* be the set of feasible allocations with no zero bundle:

$$F^* = \{x \in F \mid \forall i \in N, x_i \neq 0\}.$$

The message of the following proposition is that even with the weakest version of the Weak Pareto principle, under IIA we are not allowed much room to consider various ARSOFs.

Proposition 2 *If an ARSOF \bar{R} satisfies Weak Weak Pareto γ and IIA, then either for all $\mathbf{R} \in \mathcal{R}^n$,*

$$S_{\bar{R}}(\mathbf{R}) \subseteq \{x \in F \mid \exists i \in N, x_i = \omega\}$$

or for all $\mathbf{R} \in \mathcal{R}^n$,

$$F^* \subseteq S_{\bar{R}}(\mathbf{R}).$$

Proof. Let $\mathbf{R} \in \mathcal{R}^n$ be given. Suppose that

$$S_{\bar{R}}(\mathbf{R}) \not\subseteq \{x \in F \mid \exists i \in N, x_i = \omega\},$$

that is, there exists $y \in S_{\bar{R}}(\mathbf{R})$ such that for all $i \in N$, $y_i < \omega$. We may assume that $y \neq 0$. For if $y = 0$, then there exists $y' \in F$ such that $y' \gg 0$, and hence for all $j \in N$, $y'_j P_j y_j$. Since \bar{R} satisfies Weak Weak Pareto γ , we have $y' \in S_{\bar{R}}(\mathbf{R})$.

Thus, without loss of generality, assume that $0 < y_1 < \omega$. We need to show that $F^* \subseteq S_{\bar{R}}(\mathbf{R})$.

Step 1: We show that $\text{int}F \equiv \{x \in F \mid \forall i \in N, x_i \gg 0\} \subseteq S_{\bar{R}}(\mathbf{R})$.

Since $0 < y_1 < \omega$, there are $k, m \in \{1, \dots, \ell\}, k \neq m$ such that $y_{1k} > 0$ and $y_{1m} < \omega_m$. Without loss of generality, assume that $y_{11} > 0$ and $y_{12} < \omega_2$.

Define $z \in F$ as follows:

- (1) $z_{11} = 0$ and $z_{12} = \omega_2$,
- (2) for all $i \in N$ with $i \neq 1$, $z_{i1} = y_{i1} + \frac{y_{11}}{n-1}$ and $z_{i2} = 0$, and
- (3) for all $j \in N$ and all $k \in \{1, \dots, \ell\}$ with $k \neq 1, 2$, $z_{ik} = y_{ik}$.

Let $\mathbf{R}^0 = (R_1^0, \dots, R_n^0)$ be the profile of preferences represented by the following utility functions:

$$u_1^0(x_1) = x_{12} + \frac{1}{r_1} \sum_{m \neq 2} x_{1m},$$

$$\forall i \in N, i \neq 1, u_i^0(x_i) = x_{i1} + \frac{1}{r_i} \sum_{m \neq 1} x_{im},$$

with

$$r_1 > \frac{y_{11}}{\omega_2 - y_{12}}$$

$$\forall i \in N, i \neq 1, r_i > (n-1) \frac{y_{i2}}{y_{11}}.$$

Then, for all $j \in N$, $z_j P_j^0 y_j$. Since \bar{R} satisfies IIA, from Proposition 1, it is a constant function. Hence, $y \in S_{\bar{R}}(\mathbf{R}^0) = S_{\bar{R}}(\mathbf{R})$. Then, by Weak Weak Pareto γ , $z \in S_{\bar{R}}(\mathbf{R}^0)$.

To show that $\text{int}F \subseteq S_{\bar{R}}(\mathbf{R})$, let $t \in \text{int}F$. Let $\mathbf{R}^1 = (R_1^1, \dots, R_n^1)$ be the profile of preferences represented by the following utility functions:

$$u_1^1(x_1) = x_{11} + \frac{1}{s_1} \sum_{m \neq 1} x_{1m},$$

$$\forall i \in N, i \neq 1, u_i^1(x_i) = x_{i2} + \frac{1}{s_i} \sum_{m \neq 2} x_{im},$$

with

$$s_1 > \frac{\sum_{m \neq 1} (z_{1m} - t_{1m})}{t_{11}}$$

$$\forall i \in N, i \neq 1, s_i > \frac{\sum_{m \neq 2} (z_{im} - t_{im})}{t_{i2}}.$$

For all $j \in N$, $t_j P_j^1 z_j$. Because $z \in S_{\bar{R}}(\mathbf{R}^0)$ and $S_{\bar{R}}$ is constant, we have $z \in S_{\bar{R}}(\mathbf{R}^1)$. Then, by Weak Weak Pareto γ , $t \in S_{\bar{R}}(\mathbf{R}^1)$. Hence, $t \in S_{\bar{R}}(\mathbf{R})$.

Step 2: We show that $F^* \subseteq S_{\bar{R}}(\mathbf{R})$.

Let $y \in F^*$. Then, for all $i \in N$, $y_i \neq 0$. Let $t \in \text{int}F$ be chosen so that for each $i \in N$, there is $k(i) \in \{1, \dots, \ell\}$ such that $0 < t_{ik(i)} < y_{ik(i)}$. Let

$\mathbf{R}' = (R'_1, \dots, R'_n)$ be the profile of preferences represented by the following utility functions:

$$u_i(x_i) = x_{ik(i)} + \frac{1}{v_i} \sum_{m \neq k(i)} x_{im},$$

with

$$v_i > \frac{\sum_{m \neq k(i)} (t_{im} - y_{im})}{y_{ik(i)} - t_{ik(i)}}.$$

For all $i \in N$, $y_i P'_i t_i$. Because $t \in S_{\bar{R}}(\mathbf{R})$ and $S_{\bar{R}}$ is constant, we have $t \in S_{\bar{R}}(\mathbf{R}')$. Then, by Weak Weak Pareto γ , $y \in S_{\bar{R}}(\mathbf{R}')$. Hence, since $S_{\bar{R}}$ is constant, $y \in S_{\bar{R}}(\mathbf{R})$. \blacksquare

With Pareto-Efficiency (Weak Weak Pareto α), which is stronger than Weak Weak Pareto γ , one gets a full characterization of an ARSOF if Anonymity is introduced, as stated in the following theorem. However, the ARSOF thus characterized is special in that it selects allocations in which one individual gets all resources. Then, the second part of the theorem shows that a kind of Arrobian dictatorship is not far away. By replacing Anonymity with essential single-valuedness, one gets a dictatorial ARSOF which always gives all resources to the same individual.

Theorem 1 *There is only one ARSOF \bar{R} satisfying Pareto-Efficiency, IIA and Anonymity, namely:*

$$\forall \mathbf{R} \in \mathcal{R}^n, S_{\bar{R}}(\mathbf{R}) = \{x \in F \mid \exists i \in N, x_i = \omega\}.$$

If an ARSOF \bar{R} satisfies Pareto-Efficiency, IIA and is essentially single-valued, then

$$\exists i \in N, \forall \mathbf{R} \in \mathcal{R}^n, S_{\bar{R}}(\mathbf{R}) = \{x \in F \mid x_i = \omega\}.$$

Proof. By Proposition 2 and Pareto-Efficiency, for all $\mathbf{R} \in \mathcal{R}^n$,

$$S_{\bar{R}}(\mathbf{R}) \subseteq \{x \in F \mid \exists i \in N, x_i = \omega\}.$$

Since \bar{R} is a constant, for all $\mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$,

$$\{i \in N \mid \exists x \in S_{\bar{R}}(\mathbf{R}), x_i = \omega\} = \{i \in N \mid \exists x \in S_{\bar{R}}(\mathbf{R}'), x_i = \omega\}.$$

Therefore Anonymity requires

$$\{i \in N \mid \exists x \in S_{\bar{R}}(\mathbf{R}), x_i = \omega\} = N,$$

whereas essential single-valuedness requires

$$\{i \in N \mid \exists x \in S_{\bar{R}}(\mathbf{R}), x_i = \omega\} = \{i_0\}$$

for some fixed $i_0 \in N$. ■

From Proposition 2, the second part of Theorem 1 can be strengthened by replacing Pareto-Efficiency with Weak Weak Pareto γ in the hypothesis. It should then be noted that even with the weakest version of Weak Pareto, which does *not* require selected allocations to be Pareto-efficient, IIA and essential single-valuedness lead us to a dictatorial ARSOF.

Theorem 1 is interesting not only in its content but also in what it implies about all allocation rules of the fairness literature. Since these rules typically satisfy Pareto-Efficiency and Anonymity, and do not give all resources to one individual, *they must all violate IIA*. Proposition 1 gave the same conclusion even more immediately, since these allocation rules are not constant.

We will illustrate this important lesson with a couple of examples. In the current framework, two prominent allocation rules have been identified by the fairness literature. The first one is the **Egalitarian Walrasian AR** S_W , defined as follows: $x \in S_W(\mathbf{R})$ if $x \in F$ and there is $p \in \mathbb{R}_{++}^\ell$ such that for all $i \in N$,

$$\forall y \in \mathbb{R}_+^\ell, p \cdot y \leq p \cdot \omega/n \Rightarrow x_i R_i y.$$

Figure 1 shows how a change of preferences can alter the relative ranking of two allocations x and y without modifying individual preferences over these two allocations. In Figure 1a, x is selected and y is not, whereas the contrary obtains in Figure 1b, even though in both cases i prefers allocation y and j prefers allocation x .

The second prominent allocation rule is the **Pazner-Schmeidler AR** (see Pazner and Schmeidler 1978) S_{PS} , defined by: $x \in S_{PS}(\mathbf{R})$ if $x \in E(\mathbf{R})$ and there is $\alpha \in \mathbb{R}_+$ such that for all $i \in N$,

$$x_i I_i \alpha \omega.$$

Figure 2 is very similar to Figure 1 and illustrates the same phenomenon on the same allocations x and y , but relative to S_{PS} .

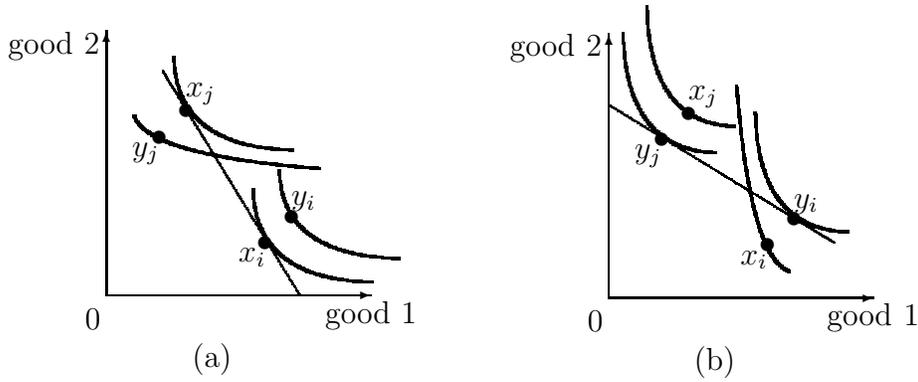


Figure 1: S_W violates IIA

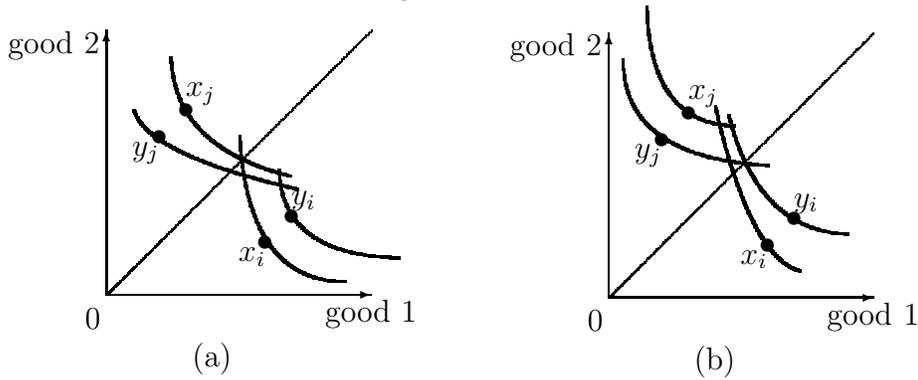


Figure 2: S_{PS} violates IIA

5 Variants of Independence of Irrelevant Alternatives

The analysis in the previous section has revealed that the success of the theory of fairness is mainly due to departure from IIA. Then, one may ask in what sense IIA is violated or, more precisely, what additional information is taken into account by ARSOFs, that is forbidden by IIA.

In the theory of social choice, the main approach with respect to information has been, following Sen (1970a, b) in particular, to introduce richer information about utilities. The theory of fairness, in contrast, has remained faithful to Arrow's initial project and usually retains only ordinal and interpersonally non-comparable information about preferences. If it introduces more information, it is about preferences, not about utilities. That is, preferences about "irrelevant" alternatives are taken into account by ARs.

It is possible to weaken IIA so as to take account of “irrelevant” alternatives (but not utilities) by strengthening the premise of the axiom in an appropriate way. This attempt brings us into several variants of the axiom, which will be introduced now. And in so doing we rely here on previous works by Hansson (1973), Fleurbaey and Maniquet (1996), and the companion paper Fleurbaey, Suzumura and Tadenuma (2001).

A first kind of additional information is contained in the marginal rates of substitution at the allocations to be compared. For efficient allocations, shadow prices enable one to compute the relative implicit income shares of different agents, thereby potentially providing a relevant measure of inequalities in the distribution of resources. Therefore, taking account of marginal rates of substitution is a natural extension of the informational basis of social choice theory in economic environments. Let $C(x_i, R_i)$ denote the cone of price vectors that support the upper contour set for R_i at x_i :

$$C(x_i, R_i) = \{p \in \mathbb{R}^\ell \mid \forall y \in \mathbb{R}_+^\ell, py = px_i \Rightarrow x_i R_i y\}.$$

When preferences R_i are strictly monotonic, one has $C(x_i, R_i) \subset \mathbb{R}_{++}^\ell$ whenever $x_i \gg 0$.

One then can require the ranking of two allocations to depend on individual preferences between these two allocations *and* also on marginal rates of substitution at these allocations, but on that only:

IIA except Marginal Rates of Substitution (IIA-MRS): $\forall x, y \in F$, $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$, if $\forall i \in N$,

$$\begin{aligned} x_i R_i y_i &\Leftrightarrow x_i R'_i y_i \\ y_i R_i x_i &\Leftrightarrow y_i R'_i x_i \\ C(x_i, R_i) &= C(x_i, R'_i) \\ C(y_i, R_i) &= C(y_i, R'_i), \end{aligned}$$

then $x \bar{R}(\mathbf{R}) y \Leftrightarrow x \bar{R}(\mathbf{R}') y$.

Marginal rates of substitution give an infinitesimally local piece of information about preferences at given allocations. A further extension of the informational basis allows the SOF to take account of finite parts of indifference hypersurfaces. The *indifference sets* are defined as

$$I(x_i, R_i) = \{z \in \mathbb{R}_+^\ell \mid z I_i x_i\}.$$

It is natural to focus on the part of indifference sets which lies within the feasible set. However, when considering any pair of allocations, the two allocations may need different amounts of total resources to be feasible and the global set F need not be relevant in its entirety. Therefore we need to introduce the following notions. The smallest amount of total resources which makes two allocations x and y feasible can be defined by $\omega(x, y) = (\omega_1(x, y), \dots, \omega_\ell(x, y))$, where for all $k \in \{1, \dots, \ell\}$:

$$\omega_k(x, y) = \max\left\{\sum_{i \in N} x_{ik}, \sum_{i \in N} y_{ik}\right\}.$$

For any vector $t \in \mathbb{R}_+^\ell$, define the set $\Omega(t) \subset \mathbb{R}_+^\ell$ by

$$\Omega(t) = \{z \in \mathbb{R}_+^\ell \mid z \leq t\}.$$

The following axiom captures the idea that the ranking of two allocations should depend only on the indifference sets, and on preferences over the minimal subset in which the two allocations are feasible.

IIA except Indifference Sets on Feasible Allocations (IIA-ISFA):
 $\forall x, y \in F, \forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$, if $\forall i \in N$,

$$\begin{aligned} I(x_i, R_i) \cap \Omega(\omega(x, y)) &= I(x_i, R'_i) \cap \Omega(\omega(x, y)) \\ I(y_i, R_i) \cap \Omega(\omega(x, y)) &= I(y_i, R'_i) \cap \Omega(\omega(x, y)), \end{aligned}$$

then $x \bar{R}(\mathbf{R}) y \Leftrightarrow x \bar{R}(\mathbf{R}') y$.

It is immediate from the definitions that

$$\begin{aligned} \text{IIA} &\Rightarrow \text{IIA-MRS} \\ &\Downarrow \\ &\text{IIA-ISFA} \end{aligned}$$

Notice that IIA-MRS does not imply IIA-ISFA because the set $I(x_i, R_i) \cap \Omega(\omega(x, y))$ does not always provide enough information to determine $C(x_i, R_i)$.⁶

It is also worthwhile here introducing a couple of independence conditions for ARs, which are closely related to IIA and its variants. Such conditions are quite common in the fairness literature. We will formulate them here for ARSOFs.

⁶It does, however, when every good is consumed by at least two agents in x .

The first one, dealing with marginal rates of substitution, is essentially Nagahisa's (1991) 'Local Independence':⁷

Independence of Preferences except MRS (IP-MRS): $\forall x \in F$, $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$, if $\forall i \in N$,

$$C(x_i, R_i) = C(x_i, R'_i),$$

then $x \in S_{\bar{R}}(\mathbf{R}) \Leftrightarrow x \in S_{\bar{R}}(\mathbf{R}')$.

The next axiom says that only the parts of indifference sets concerning feasible allocations should matter.

Independence of Preferences except Indifference Sets on Feasible Allocations (IP-ISFA): $\forall x \in F$, $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$, if $\forall i \in N$,

$$I(x_i, R_i) \cap \Omega(\omega) = I(x_i, R'_i) \cap \Omega(\omega),$$

then $x \in S_{\bar{R}}(\mathbf{R}) \Leftrightarrow x \in S_{\bar{R}}(\mathbf{R}')$.

Although these independence conditions may seem restrictive, they are actually not really stronger than the previous IIA axioms.

Proposition 3 *On the class of ARSOFs that never select $0 = (0, \dots, 0)$, IIA-MRS \Rightarrow IP-MRS, and IIA-ISFA \Rightarrow IP-ISFA.*

Proof. IIA-MRS \Rightarrow IP-MRS. Let $x \in S_{\bar{R}}(\mathbf{R})$ and \mathbf{R}' be such that for all $i \in N$, $C(x_i, R'_i) = C(x_i, R_i)$. Notice that $0 = (0, \dots, 0) \notin S_{\bar{R}}(\mathbf{R})$. Since for all $i \in N$, $C(0, R'_i) = C(0, R_i) = \mathbb{R}_+^\ell$, and $x_i R_i 0 \Leftrightarrow x_i R'_i 0$, and $0 R_i x_i \Leftrightarrow 0 R'_i x_i$, it follows from IIA-MRS that $x \in S_{\bar{R}}(\mathbf{R}')$ and $0 \notin S_{\bar{R}}(\mathbf{R}')$.

IIA-ISFA \Rightarrow IP-ISFA. Let $x \in S_{\bar{R}}(\mathbf{R})$ and \mathbf{R}' be such that for all $i \in N$, $I(x_i, R_i) \cap \Omega(\omega) = I(x_i, R'_i) \cap \Omega(\omega)$. Notice that for all $i \in N$, $I(0, R'_i) = I(0, R_i) = \{0\}$. Then, by IIA-ISFA, $x \in S_{\bar{R}}(\mathbf{R}')$. ■

It is also easy to check that IP-MRS implies IIA-MRS, and that, for ARSOFs which never select allocations x such that $\sum_{i \in N} x_i \neq \omega$, IP-ISFA implies IIA-ISFA. In other words, for all practical purposes, the distinction between the IP axioms introduced here and their IIA counterparts is negligible.

⁷See also Yoshihara (1998).

6 Informational requirements for allocation rules

Even though the allocation rule characterized in Theorem 1 above is fully anonymous, it is not appealing because it contains only strongly unequal allocations. A minimal requirement of equality is the following:

Equal Treatment of Equals (for ARSOFs): $\forall \mathbf{R} \in \mathcal{R}^n, \forall x \in S_{\bar{R}}(\mathbf{R}), \forall i, j \in N, \text{ if } R_i = R_j, \text{ then } x_i I_i x_j.$

This requirement is very minimal, and one may notice that any ARSOF \bar{R} satisfying Anonymity and essential single-valuedness necessarily satisfies Equal Treatment of Equals.

From the statement and the proof of Theorem 1 we immediately deduce:

Corollary 1 *There is no ARSOF satisfying Pareto-Efficiency, IIA and Equal Treatment of Equals.⁸ There is no essentially single-valued ARSOF satisfying Pareto-Efficiency, IIA and Anonymity.*

The question we ask in this section is how much IIA needs to be weakened, or how much additional information is needed in order to obtain the existence of an ARSOF satisfying the above sets of conditions.

Our first result is that with IIA-MRS, a possibility is obtained, but there remains a difficulty about essential single-valuedness.

Theorem 2 *There exists an ARSOF satisfying Pareto-Efficiency, IIA-MRS, Equal Treatment of Equals and Anonymity. There is no essentially single-valued ARSOF satisfying Pareto-Efficiency, IIA-MRS and Equal Treatment of Equals. There is no essentially single-valued ARSOF satisfying Pareto-Efficiency, IIA-MRS and Anonymity.*

Proof. The possibility is illustrated by the Egalitarian Walrasian ARSOF \bar{R}_{SW} .

The second impossibility is implied by the first impossibility because essential single-valuedness and Anonymity imply Equal Treatment of Equals. To show the first impossibility, suppose, to the contrary, that there exists

⁸A slightly different proof obtains by showing that the only constant ARSOF satisfying Equal Treatment of Equals selects the egalitarian allocation giving ω/n to every agent, which is not Pareto-efficient in general.

an essentially single-valued ARSOF \bar{R} satisfying Pareto-Efficiency, IIA-MRS and Equal Treatment of Equals. By Pareto-Efficiency, for all $\mathbf{R} \in \mathcal{R}^n$, $0 = (0, \dots, 0) \notin S_{\bar{R}}(\mathbf{R})$. Hence, from Proposition 3, \bar{R} satisfies IP-MRS.

Let \mathcal{R}^* be the subset of \mathcal{R} such any $R \in \mathcal{R}^*$ is representable by a utility function of the following kind:

$$u(x_1, \dots, x_\ell) = f_1(x_1) + \dots + f_\ell(x_\ell),$$

where for all $k \in \{1, \dots, \ell\}$, f_k is continuous, increasing, concave, and differentiable over \mathbb{R}_{++} , with $\lim_{x \rightarrow 0} f'_k(x) = +\infty$. The relevant property of this domain is that for all $\mathbf{R} \in (\mathcal{R}^*)^n$,

$$E(\mathbf{R}) \subseteq \{x \in \mathbb{R}_+^\ell \mid \forall i \in N, x_i \gg 0 \text{ or } x_i = 0\}.$$

Let $\mathbf{R} \in (\mathcal{R}^*)^n$ be given.

Firstly, suppose that there is $x \in S_{\bar{R}}(\mathbf{R}) \setminus S_W(\mathbf{R})$. By Pareto-Efficiency $x \in E(\mathbf{R})$. Hence, we have $x_i \gg 0$ or $x_i = 0$ for all $i \in N$, and by differentiability of preferences there is a shadow price vector $p \in \mathbb{R}_{++}^\ell$ such that

$$\forall i \in N, C(x_i, R_i) = \{\lambda p \mid \lambda \in \mathbb{R}_{++}\} \text{ or } x_i = 0.$$

For this p , define $R^p \in \mathcal{R}$ by

$$\forall z, z' \in \mathbb{R}_+^\ell, z R^p z' \Leftrightarrow p \cdot z \geq p \cdot z'.$$

Let $\mathbf{R}^p = (R^p, \dots, R^p) \in \mathcal{R}^n$. By IP-MRS, $x \in S_{\bar{R}}(\mathbf{R}^p)$. Since $x \notin S_W(\mathbf{R})$, there exist $i, j \in N$ such that $x_i P^p x_j$, in contradiction to Equal Treatment of Equals. As a consequence, $S_{\bar{R}}(\mathbf{R}) \subset S_W(\mathbf{R})$.

Secondly, suppose that there is $x \in S_W(\mathbf{R}) \setminus S_{\bar{R}}(\mathbf{R})$. For all $i \in N$, let $\mathbf{R}' \in (\mathcal{R}^*)^n$ be a profile of homothetic (a given R in \mathcal{R}^* is homothetic if all its component functions f_k are homogeneous of the same degree) and strictly convex preferences satisfying

$$\forall i \in N, C(x_i, R'_i) = C(x_i, R_i).$$

We have $x \in S_W(\mathbf{R}')$. Moreover, by Theorem 1 in Eisenberg (1961), all allocations in $S_W(\mathbf{R}')$ are Pareto-indifferent. By strict convexity of preferences, one therefore has $S_W(\mathbf{R}') = \{x\}$. Since, by the previous argument, $S_{\bar{R}}(\mathbf{R}') \subset S_W(\mathbf{R}')$, we have $S_{\bar{R}}(\mathbf{R}') = \{x\}$. By IP-MRS, $x \in S_{\bar{R}}(\mathbf{R})$, which is a contradiction. Therefore $S_W(\mathbf{R}) \subset S_{\bar{R}}(\mathbf{R})$.

In conclusion, $S_{\bar{R}}(\mathbf{R}) = S_W(\mathbf{R})$ for all $\mathbf{R} \in (\mathcal{R}^*)^n$. But S_W is not essentially single-valued on the whole domain $(\mathcal{R}^*)^n$. This contradicts essential single-valuedness of $S_{\bar{R}}$. ■

Only with IIA-ISFA do we really obtain a full possibility result.

Theorem 3 *There exists an essentially single-valued ARSOF satisfying Pareto-Efficiency, IIA-ISFA, Anonymity and Equal Treatment of Equals.*

Proof. Consider the Pazner-Schmeidler ARSOF $\bar{R}_{S_{PS}}$, defined at the end of section 4. It obviously satisfies Pareto-Efficiency, Anonymity and Equal Treatment of Equals. To check that it satisfies IIA-ISFA, let $x, y \in F$ and $\mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$ be such that for all $i \in N$,

$$\begin{aligned} I(x_i, R_i) \cap \Omega(\omega(x, y)) &= I(x_i, R'_i) \cap \Omega(\omega(x, y)) \\ I(y_i, R_i) \cap \Omega(\omega(x, y)) &= I(y_i, R'_i) \cap \Omega(\omega(x, y)), \end{aligned}$$

and $x \in S_{PS}(\mathbf{R})$ and $y \notin S_{PS}(\mathbf{R})$. Let $\alpha \in \mathbb{R}_+$ be such that for all $i \in N$, $x_i I_i \alpha\omega$. Then, necessarily $\alpha < 1$. Notice that $\sum_{i \in N} x_i = \omega$ because $x \in E(\mathbf{R})$. Hence, $\Omega(\omega(x, y)) = \Omega(\omega)$, and $\alpha\omega \in \Omega(\omega(x, y))$. Together with the above equalities, we deduce that $x \in S_{PS}(\mathbf{R}')$ and $y \notin S_{PS}(\mathbf{R}')$. ■

7 Under Weak Pareto, social ordering functions need more information

Fleurbaey and Maniquet (1996), in this model, showed that there exist many SOFs satisfying Weak Pareto, Anonymity and the following weak version of IIA:

IIA except Whole Indifference Sets (IIA-WIS): $\forall x, y \in F, \forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n$, if $\forall i \in N$,

$$\begin{aligned} I(x_i, R_i) &= I(x_i, R'_i) \\ I(y_i, R_i) &= I(y_i, R'_i), \end{aligned}$$

then $x \bar{R}(\mathbf{R}) y \Leftrightarrow x \bar{R}(\mathbf{R}') y$.

This axiom is weaker than all IIA axioms considered above, and one may ask what is the minimal amount of information needed by a SOF in order

to satisfy Weak Pareto and Anonymity (or Non-Dictatorship). In Fleurbaey, Suzumura and Tadenuma (2001), we showed that no SOF \bar{R} satisfies Weak Pareto, Non-Dictatorship (over the subset X of allocations in which no agent has a zero bundle), and either IIA-MRS or IIA-ISFA.

But these results were obtained in the particular case of unbounded resources $F = \mathbb{R}_+^{n\ell}$. The bounded case on which we focus here has attracted less attention in the social choice literature,⁹ and here we have the following result.

Theorem 4 *There is no SOF \bar{R} satisfying Weak Pareto, IIA-MRS and Anonymity. There is no SOF \bar{R} satisfying Weak Pareto, IIA-ISFA and Anonymity.*

Proof. In order to prove the impossibility, it is convenient to consider different possible sizes of the population. For each $a \in \mathbb{R}_+^\ell$, define

$$B(a) = \{b \in \mathbb{R}_+^\ell \mid \max_{k \in \{1, \dots, \ell\}} |b_k - a_k| \leq \frac{1}{10}\}$$

Case $n = 2$. Consider the bundles $x = (8, 1/2, 0, \dots)$, $y = (12, 1/2, 0, \dots)$, $z = (1/2, 12, 0, \dots)$, $w = (1/2, 8, 0, \dots)$. Let preferences R_1 and R_2 be defined as follows. On the subset

$$S_1 = \{v \in \mathbb{R}_+^\ell \mid \forall i \in \{3, \dots, \ell\}, v_i = 0 \text{ and } v_2 \leq \min\{v_1, 1\}\}$$

one has

$$vR_1v' \Leftrightarrow v_1 + 2v_2 \geq v'_1 + 2v'_2,$$

and on the subset

$$S_2 = \{v \in \mathbb{R}_+^\ell \mid \forall i \in \{3, \dots, \ell\}, v_i = 0 \text{ and } v_1 \leq \min\{v_2, 1\}\},$$

one has

$$vR_1v' \Leftrightarrow 2v_1 + v_2 \geq 2v'_1 + v'_2.$$

On $B(x) \cup B(y)$, one has

$$vR_1v' \Leftrightarrow v_1 + 2v_2 + \sum_{k=3}^{\ell} v_k \geq v'_1 + 2v'_2 + \sum_{k=3}^{\ell} v'_k,$$

⁹An exception is Bordes, Campbell and Le Breton (1995), in which Arrow's theorem is extended to the case of a bounded set of allocations.

and on $B(z) \cup B(w)$,

$$vR_1v' \Leftrightarrow 2v_1 + v_2 + \sum_{k=3}^{\ell} v_k \geq 2v'_1 + v'_2 + \sum_{k=3}^{\ell} v'_k.$$

Since

$$w_1 + (1 - w_1) + 2[w_2 - 2(1 - w_1)] > x_1 + 2x_2$$

and

$$2[y_1 - 2(1 - y_2)] + y_2 + (1 - y_2) > 2z_1 + z_2,$$

it is possible to complete the definition of R_1 such that wP_1x and yP_1z . Then define R_2 so that it coincides with R_1 on $S_1 \cup S_2$, and on $B(a)$ for all $a \in \{x, y, z, w\}$. Similarly, it is possible to complete the definition of R_2 such that xP_2w and zP_2y . Figure 3 illustrates this construction.

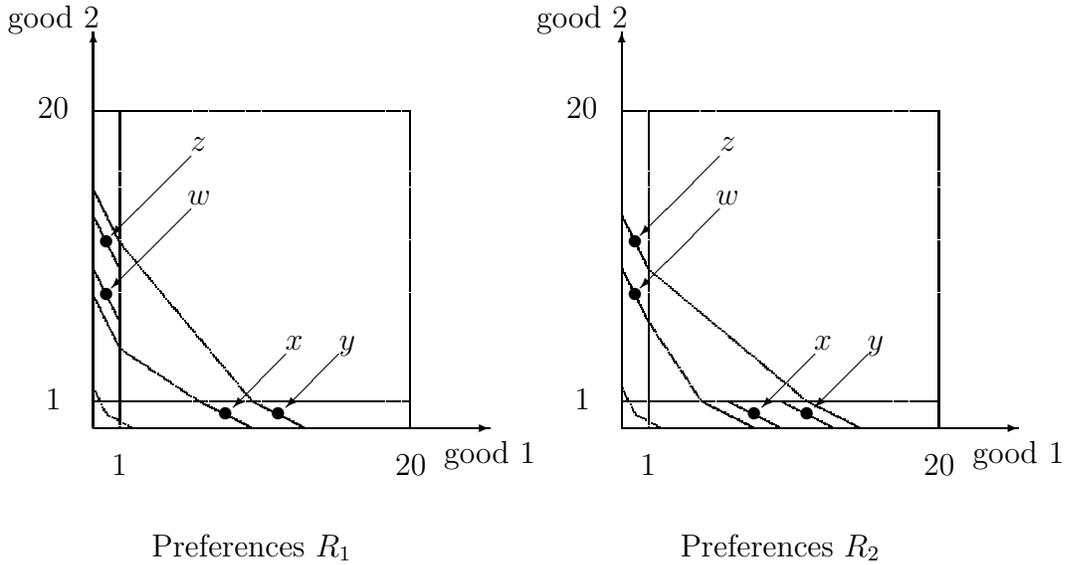


Figure 3: Construction of R_1 and R_2

If the profile of preferences is $\mathbf{R} = (R_1, R_2)$, by Weak Pareto one has:

$$(y, x)\bar{P}(\mathbf{R})(z, w) \text{ and } (w, z)\bar{P}(\mathbf{R})(x, y).$$

If the profile of preferences is $\mathbf{R}' = (R_1, R_1)$, by Anonymity one has:

$$(y, x)\bar{I}(\mathbf{R}')(x, y) \text{ and } (w, z)\bar{I}(\mathbf{R}')(z, w).$$

Since R_1 and R_2 coincide on $S_1 \cup S_2$, and on $B(a)$ for all $a \in \{x, y, z, w\}$, by IIA-MRS or IIA-ISFA, one has:

$$(y, x)\bar{I}(\mathbf{R}')(x, y) \Leftrightarrow (y, x)\bar{I}(\mathbf{R})(x, y)$$

$$\text{and } (w, z)\bar{I}(\mathbf{R}')(z, w) \Leftrightarrow (w, z)\bar{I}(\mathbf{R})(z, w).$$

By transitivity, one gets $(x, y)\bar{P}(\mathbf{R})(x, y)$, which is impossible.

Case $n = 3$. Consider the bundles $x = (8, 1/3, 0, \dots)$, $y = (12, 1/3, 0, \dots)$, $t = (10, 1/3, 0, \dots)$, $z = (1/3, 12, 0, \dots)$, $w = (1/3, 8, 0, \dots)$, $r = (1/3, 10, 0, \dots)$. Let preferences R_1 , R_2 and R_3 be defined as above on the subset $S_1 \cup S_2$, and on $B(a)$ for all $a \in \{x, y, z, w\}$. Complete their definition so that yP_1z , wP_1x , tP_2r , zP_2y , xP_3w , rP_3t .

If the profile of preferences is $\mathbf{R} = (R_1, R_2, R_3)$, by Weak Pareto one has:

$$(y, t, x)\bar{P}(\mathbf{R})(z, r, w) \text{ and } (w, z, r)\bar{P}(\mathbf{R})(x, y, t).$$

If the profile of preferences is $\mathbf{R}' = (R_1, R_1, R_1)$, by Anonymity one has:

$$(y, t, x)\bar{I}(\mathbf{R}')(x, y, t) \text{ and } (w, z, r)\bar{I}(\mathbf{R}')(z, r, w).$$

Since R_1, R_2 and R_3 coincide on $S_1 \cup S_2$, and on $B(a)$ for all $a \in \{x, y, z, w\}$, by IIA-MRS or IIA-ISFA, one has:

$$(y, t, x)\bar{I}(\mathbf{R}')(x, y, t) \Leftrightarrow (y, t, x)\bar{I}(\mathbf{R})(x, y, t)$$

$$\text{and } (w, z, r)\bar{I}(\mathbf{R}')(z, r, w) \Leftrightarrow (w, z, r)\bar{I}(\mathbf{R})(z, r, w).$$

By transitivity, one gets $(x, y, t)\bar{P}(\mathbf{R})(x, y, t)$, which is impossible.

Case $n = 2k$. Partition the population into k pairs, and construct an argument similar to the case $n = 2$, with the bundles $x = (8, 1/n, 0, \dots)$, $y = (12, 1/n, 0, \dots)$, $z = (1/n, 12, 0, \dots)$, $w = (1/n, 8, 0, \dots)$, and the allocations (y, x, y, x, \dots) , (x, y, x, y, \dots) , (z, w, z, w, \dots) and (w, z, w, z, \dots) .

Case $n = 2k + 1$. Partition the population into $k - 1$ pairs and one triple, and construct an argument combining the cases $n = 2$ and $n = 3$, with the bundles $x = (8, 1/n, 0, \dots)$, $y = (12, 1/n, 0, \dots)$, $t = (10, 1/n, 0, \dots)$, $z = (1/n, 12, 0, \dots)$, $w = (1/n, 8, 0, \dots)$, $r = (1/n, 10, 0, \dots)$, and the allocations $(y, x, y, x, \dots, y, t, x)$, $(x, y, x, y, \dots, x, y, t)$, $(z, w, z, w, \dots, z, r, w)$ and $(w, z, w, z, \dots, w, z, r)$. ■

This result proves that under Weak Pareto, more information about preferences is needed than under Pareto-Efficiency. In that sense, it is true that

the theory of fairness, with its coarse orderings, is less demanding in information than the theory of social choice.

As explained in Fleurbaey, Suzumura and Tadenuma (2001), however, one should not conclude from this analysis that full knowledge of indifference curves is needed under Weak Pareto. Define the Pazner-Schmeidler SOF \bar{R}_{PS} as follows: $x \bar{R}(\mathbf{R}) y$ if and only if

$$\min\{\alpha \in \mathbb{R}_+ \mid \exists i \in N, \alpha \omega R_i x_i\} \geq \min\{\alpha \in \mathbb{R}_+ \mid \exists i \in N, \alpha \omega R_i y_i\}.$$

This SOF satisfies Weak Pareto and Anonymity, even though it only requires knowledge of the intersection of indifference curves with a ray from the origin. In addition, although this SOF does not satisfy IIA-ISFA in the current framework, it can be shown to satisfy IIA-ISFA when only allocations of the subset

$$\{x \in \mathbb{R}_+^{n\ell} \mid \sum_{i \in N} x_i = \omega\}$$

with no free disposal, instead of F , are ranked.

8 Toward a unified theory

There have been many attempts to import fairness concepts into social choice, and thereby build a unified theory, such as Feldman and Kirman (1974), Varian (1976), Suzumura (1981a,b, 1983) and Tadenuma (2002). But they did not focus on the informational requirements to obtain positive results.

Our approach provides a unified framework which covers the theory of social choice and the theory of fairness. Because ARs in the theory of fairness are isomorphic to ARSOFs in the theory of social choice, and ARSOFs are just a particular kind of SOF, the concept of SOF is comprehensive enough to encompass all relevant notions. This shows how the theory of fairness is, rigorously, a part of the theory of social choice.

As a consequence, the way in which possibility results are obtained with ARs, by broadening the informational basis, can be adopted for SOFs, albeit, as shown above, the amount of additional information needed is greater. From this perspective, there is no longer any reason to view the theory of social choice as plagued with impossibilities, and no longer any reason for social choice theorists to envy fairness theorists and their positive results.

The same recipe for success can be adopted by social choice theorists.¹⁰

In this section we examine two possible objections to this proposed integration of fair allocation theory into social choice theory. The first objection would go by recollecting that the celebrated Arrow Program of social choice theory consists of two separate steps, viz., (a) the construction of a social preference ordering corresponding to each and every profile of individual preference orderings; and (b) the construction of a social choice function in terms of the optimization of social preferences within each and every set of feasible social alternatives. The first step, which may be called the preference aggregation stage, is meant to determine the uniform social objective before the set of feasible social alternatives is revealed. The second step, which may be called the social choice stage, is meant to determine the rational social choice after the set of feasible social alternatives is revealed. Even though in the theory of fair allocation we may construct a coarse social ordering in terms of the fair allocations versus unfair allocations, such an ordering hinges squarely on the specification of the set of feasible allocations. Thus, the objection goes, in view of the basic scenario of the Arrow Program of social choice theory, the theory of fair allocation does not really offer much to the preference aggregation stage of social choice theory.

Our response to this objection is that what is called “social choice theory” in this paper actually encompasses the preference aggregation stage of the Arrow program, as presented above, as a particular case. We believe that it is quite convenient to see the common formal structure in all exercises of construction of a preference ordering over a set of alternatives, whether this set is determined by feasibility constraints or not. In this paper, the need to compare the social choice approach and the fair allocation approach has led us to retain

$$F = \{x \in \mathbb{R}_+^\ell \mid x_1 + \dots + x_n \leq \omega\}$$

as the relevant set of alternatives. An orthodox vision of the Arrow Program of social choice theory might possibly require the construction of the social preference ordering to be made on the full set $\mathbb{R}_+^{n\ell}$, rather than F , but we do not think that the construction of a social preference ordering over F should be excluded from social choice theory for that reason.¹¹ Moreover, the

¹⁰For characterizations of SOFs based on fairness axioms, see e.g. Fleurbaey and Maniquet (2000, 2001).

¹¹Arrow himself was actually vague about the set of alternatives in his monograph on social choice. For instance, in the economic example he introduces in chap. 6, sect. 4, he

notion of feasibility itself is multi-faceted. Although F is determined by some feasibility constraints, the set of actually feasible alternatives, in practical applications, is likely to be a strict subset of F . For instance, the political system may give special value to a status quo x_0 , and restrict attention to another particular alternative x , introduced as a proposed reform of the status quo. In order to decide whether x is better than x_0 or not, a fine-grained ranking of all members of F is quite useful, and a ranking of all members of $\mathbb{R}_+^{n\ell}$ would be perfectly adequate as well, but would be more than needed.

The second objection to our unification would rely on an alternative way of unifying the two theories, which has been elegantly formulated in Fishburn (1973) and adapted to economic environments by Le Breton (1997). It consists in broadening the concept of AR, as done in the theory of social choice based on social decision rules (SDR).

Let \mathcal{F} denote the set of non-empty subsets of F , and let $\mathcal{A} \subset \mathcal{F}$. An SDR is a mapping \bar{S} from $\mathcal{R}^n \times \mathcal{A}$ to \mathcal{F} such that for all $\mathbf{R} \in \mathcal{R}^n$, all $A \in \mathcal{A}$, $\bar{S}(\mathbf{R}, A) \subset A$ and $\bar{S}(\mathbf{R}, A) \neq \emptyset$. The subset A is called an agenda, and \mathcal{A} is the class of agendas.

An AR is just a particular kind of SDR, for which $\mathcal{A} = \{F\}$. And one can recover a SOF from an SDR if \mathcal{A} contains all pairs of allocations $\{x, y\} \subset F$ and satisfies a choice consistency condition. The derived SOF $\bar{R}_{\bar{S}}$ is then defined by:

$$x \bar{R}_{\bar{S}}(\mathbf{R}) y \Leftrightarrow x \in \bar{S}(\mathbf{R}, \{x, y\}).$$

In this perspective, the specificity of the theory of fairness is that it has a very restricted class of agendas. This expresses the fact that the theory of fairness only seeks the good allocations among all feasible ones, whereas the theory of social choice wants to make fine-grained selections in most conceivable agendas.

That possibility results are obtained in the theory of fairness is likely to be interpreted, in this approach, as due to the restricted agendas, and this reinforces the usual explanation which opposes fine-grained social preferences and selection. But this would be a hasty conclusion. Arrow's independence condition, applied to SDRs, is formulated as follows in Le Breton (1997):

simply states: 'Suppose that among the *possible* alternatives there are three, none of which gives any individual at least as much of both commodities as any other.' (Arrow 1963, p. 68; emphasis added) Bordes, Campbell and Le Breton (1995) study Arrow's theorem on F as a relevant social choice exercise.

Independence of Infeasible Alternatives (IIF): $\forall \mathbf{R}, \mathbf{R}' \in \mathcal{R}^n, \forall A \in \mathcal{A}$,
if $\forall i \in N$,

$$\forall x, y \in A : x_i R_i y_i \Leftrightarrow x_i R'_i y_i,$$

then $\bar{S}(\mathbf{R}, A) = \bar{S}(\mathbf{R}', A)$.

When the class of agendas is restricted, the amount of information about preferences that may be used by \bar{S} when considering to choose x as against y increases automatically, because the subset A on which preference information is retained is larger. Therefore going to restricted agendas has two consequences. First, it makes one go from fine-grained preferences to coarse preferences, as emphasized by the usual explanation of the possibility results in fairness theory. Second, and, we believe, more importantly, it increases the amount of relevant information about preferences, as delineated by IIF.

Although the two unified theories (in terms of SOFs or in terms of SDRs) are essentially isomorphic, we are inclined to think that the SOF approach developed in this paper is more suitable to the analysis of the informational basis of the various theories. For instance, consider the Egalitarian Walrasian AR which only needs knowledge of marginal rates of substitution to decide whether an allocation is selected or not. In the SOF approach, this is easily captured by the IIA-MRS axiom, which is clearly a weakening of IIA. In the SDR approach, there is no as easy a way to modify IIF in order to capture the same idea. Because in order to check that $\bar{S}(\mathbf{R}, F) = \bar{S}(\mathbf{R}', F)$, that is, $S_W(\mathbf{R}) = S_W(\mathbf{R}')$, one needs to know marginal rates of substitutions at many allocations, and this requires global knowledge of \mathbf{R} over a large subset of allocations. Moreover, S_W , viewed as an SDR, does not even satisfy IIF because at corner allocations marginal rates of substitution may depend on preferences outside F .

9 Conclusion

In the traditional theory of social choice in economic environments, Fleurbaey, Suzumura and Tadenuma (2001) have shown that the construction of an Arrovian social ordering function, in a framework with purely ordinal, non-comparable preferences, requires information about the shape of indifference curves that goes well beyond purely local data such as marginal rates of substitution.

The main lesson of this paper is that even for the less ambitious project

of constructing allocation rules, it is also necessary to introduce more information than allowed by the Arrovian independence of irrelevant alternatives. And the second lesson is that, nonetheless, a purely local information such as marginal rates of substitution is sufficient (or almost so) for allocation rules, whereas it is not so for social ordering functions.

We hope that our paper, more broadly, contributes to clarifying the informational foundations in the theory of social choice and in the theory of fair allocation, and also to clarifying the links and differences between these two theories. Our proposal for a unified theory of social choice, where possibility results from the fairness part can be extended to SOFs, should shake off the negative fame of social choice.

10 References

- ARROW K. J. 1950, "A Difficulty in the Concept of Social Welfare", *Journal of Political Economy* 58: 328-346.
- ARROW K. J. 1951, *Social Choice and Individual Values*, New York: Wiley. Second ed., 1963.
- BORDES G., D. E. CAMPBELL, M. LE BRETON 1995, "Arrow's Theorems for Economic Domains and Edgeworth Hyperboxes", *International Economic Review* 36: 441-454.
- EISENBERG E. 1961, "Aggregation of Utility Functions", *Management Science* 7(4): 337-350.
- FELDMAN A. M., A. KIRMAN 1974, "Fairness and Envy", *American Economic Review* 64: 996-1005.
- FISHBURN P. 1973, *The Theory of Social Choice*, Princeton: Princeton University Press.
- FLEURBAEY M., F. MANIQUET 1996, "Utilitarianism versus Fairness in Welfare Economics", forthcoming in M. Salles and J. A. Weymark (Eds), *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls*, Cambridge : Cambridge University Press.
- FLEURBAEY M., F. MANIQUET 2000, "Fair Social Orderings with Unequal Production Skills", mimeo, U. of Cergy.
- FLEURBAEY M., F. MANIQUET 2001, "Fair Social Orderings", mimeo, U. of Pau.
- FLEURBAEY M., K. SUZUMURA, K. TADENUMA 2001, "Arrovian Aggregation in Economic Environments: How Much Should We Know About Indif-

- ference Surfaces?" mimeo., Hitotsubashi University and U. of Pau.
- FOLEY D. 1967, "Resource Allocation and the Public Sector", *Yale Economic Essays* 7: 45-98.
- HANSSON B. 1973, "The Independence Condition in the Theory of Social Choice", *Theory and Decision* 4: 25-49.
- KOLM S. C. 1972, *Justice et Équité*, Paris: Ed. du CNRS.
- KOLM S. C. 1996, *Modern Theories of Justice*, Cambridge, Mass.: MIT Press.
- LE BRETON M. 1997, "Arrowian Social Choice on Economic Domains", in K. J. Arrow, A. Sen, K. Suzumura (Eds), *Social Choice Re-examined*, Vol. 1, London: Macmillan and New York: St. Martin's Press.
- MOULIN H., W. THOMSON 1997, "Axiomatic Analysis of Resource Allocation Problems", in K. J. Arrow, A. Sen, K. Suzumura (eds.), *Social Choice Re-examined*, Vol. 1, London: Macmillan and New York: St. Martin's Press.
- NAGAHISA R.-I. 1991, "A Local Independence Condition for Characterization of Walrasian Allocation Rule", *Journal of Economic Theory* 54: 106-123.
- NAGAHISA R.-I., S. C. SUH 1995, "A Characterization of the Walras Rule", *Social Choice and Welfare* 12: 335-352.
- PAZNER E., D. SCHMEIDLER 1978, "Egalitarian-Equivalent Allocations: A New Concept of Economic Equity", *Quarterly Journal of Economics* 92: 671-687.
- ROBERTS K. 1980, "Interpersonal Comparability and Social Choice Theory", *Review of Economic Studies* 47: 421-439.
- ROEMER J. E. 1996, *Theories of Distributive Justice*, Cambridge: Harvard U. Press.
- SEN A. K. 1970a, "Interpersonal Aggregation and Partial Comparability", *Econometrica* 38: 393-409.
- SEN A. K. 1970b, *Collective Choice and Social Welfare*, San-Francisco: Holden-Day; republished Amsterdam: North-Holland, 1979.
- SEN A. K. 1986, "Social Choice Theory", in K. J. Arrow, M. D. Intriligator (eds.), *Handbook of Mathematical Economics*, Vol. 3, Amsterdam: North-Holland.
- SUZUMURA K. 1980, "Liberal Paradox and the Voluntary Exchange of Rights-Exercising", *Journal of Economic Theory* 22: 407-422.
- SUZUMURA K. 1981a, "On Pareto-Efficiency and the No-Envy Concept of Equity", *Journal of Economic Theory* 25: 367-379.
- SUZUMURA K. 1981b, "On the Possibility of 'Fair' Collective Choice Rule", *International Economic Review* 22: 351-364.

- SUZUMURA K. 1983, *Rational Choice, Collective Decisions, and Social Welfare*, Cambridge: Cambridge U. Press.
- TADENUMA K. 2002, "Efficiency First or Equity First? Two Principles and Rationality of Social Choice", *Journal of Economic Theory* 104: 462-472.
- VARIAN H. 1974, "Equity, Envy, and Efficiency", *Journal of Economic Theory* 9: 63-91.
- VARIAN H. 1976, "Two Problems in the Theory of Fairness", *Journal of Public Economics* 5: 249-260.
- WEYMARK J. 1984, "Arrow's Theorem with Social Quasi-Orderings", *Public Choice* 42: 235-246.
- YOSHIHARA N. 1998, "Characterizations of the Public and Private Ownership Solutions", *Mathematical Social Sciences* 35: 165-184.