PART I

BASICS

This first part of this book is an introduction to the approach that is developed. Its first aim is to help the reader understand where the approach lies in the broad field of normative economics; the first chapter reviews the various subdomains of the field to highlight the main differences. Its main point is to show that our approach is the only one that combines various features that are scattered in various classical approaches. Fundamentally, this approach opens a working space at the intersection of social choice theory and fair allocation theory. Like the former, it constructs rankings of all possible alternatives; like the latter, it involves fairness principles about resource allocation rather than interpersonal comparisons of utility; like both theories, it puts the Pareto principle, the ideal of respecting individual preferences, first in the order of priorities. The expression "fair social choice" is often used as a name for the approach.[1]

The second and third chapters introduce general results that appear to be common to all models that have been studied so far. These results are striking and somewhat counterintuitive. The first is that there is a tension between the Pareto principle and the deceptively simple idea that an agent who has more in all dimensions than another, such as one who consumes more of all commodities, is necessarily better off and should transfer some of the surplus to the other agent. The idea that "having more in all dimensions" is an obvious situation of advantage has been flagged by Sen (1985, 1992) as a partial but robust solution to the problem of indexing well-being in a multidimensional context. This seems indeed very natural, but our approach, because it involves respecting individual preferences, implies that it must sometimes happen that the better-off agent is the one who has less in all dimensions. As we will explain, this is in fact much less counterintuitive than it seems when individual preferences are taken into account.

The second result is that, once again because of the multidimensional context and the respect of individual preferences, a minimal degree of inequality aversion in the social criterion implies that one must actually give absolute priority to the worst-off. This is surprising because it appears easy to use individual

[1] It is more transparent but less entertaining than the alternative name "welfair economics."

utility functions with a certain degree of concavity to obtain a finite preference, in the social ordering, for equality in resources. As it turns out, it is not easy at all. The construction of such utility functions is so difficult and informationally demanding that the only "simple" approach consists of adopting the absolute priority for the worst-off. As illustrated several times in this book, this does not mean that this approach is only for radicals. Giving absolute priority for the worst-off is compatible with many possible ways of identifying the worst-off. As we will show, even free-market libertarians can see their ideas reflected in particular social criteria developed along this vein.

The fourth and last chapter of this part may be skipped by the reader who is more interested in applications than in theoretical underpinnings. It examines the informational requirements of our approach. In the theory of social choice, following Sen (1970), it has become classical to analyze the problem of finding possibility results as a problem about the information that is used in the construction of social preferences. This is justified, and in Chapter 4 we show how our possibility results are linked to the fact that our social preferences involve certain kinds of interpersonal comparisons. However, they do not perform interpersonal comparisons of utility but instead compare resource bundles or, more precisely, indifference curves (or sets, in more than two dimensions). We also examine in that chapter other aspects of information that are important in understanding the approach, such as the fact that social preferences may depend in a limited way on the feasible set.

Cambridge University Press
978-0-521-71534-8 - A Theory of Fairness and Social Welfare
Marc Fleurbaey and Francois Maniquet
Excerpt
More information

CHAPTER 1

A Contribution to Welfare Economics

1.1 INTRODUCTION

Social welfare and fairness are concepts with a venerable history in economic theory. In this chapter, we situate the approach that is developed in this book within the field of welfare economics. The simplest way to do this is to compare it with the various existing approaches in the field. For each of them, we list the main features that are shared with our approach, and explain why we keep them. We also present the main differences and explain and justify why we have to, or choose to, depart from those classical approaches. We hope that this short overview clarifies how our undertaking can contribute to the development of some of these subfields.

This discussion relies on a simple example. Assume that a positive quantity of several divisible private goods must be distributed to a population of agents, each of whom has personal preferences over his or her own consumption.[1] We discuss each classical subfield of welfare economics, as well as our approach, in this simple framework.

The preferences are assumed to be well-behaved and self-centered (i.e., without consumption externalities). What we call an *economy* is a population with a profile of preferences and a social endowment to be distributed. Formally, we consider that there are ℓ goods (with $\ell \geq 2$). The social endowment of goods is denoted $\Omega \in \mathbb{R}^{\ell}_{++}$. The population is a nonempty finite set N.

Each agent i in N has a preference relation R_i, which is a complete ordering over bundles z_i belonging to agent i's consumption set $X = \mathbb{R}^{\ell}_{+}$. For two bundles $x, y \in X$, we write $x \; R_i \; y$ to denote that agent i is at least as well off at x as at y. The corresponding strict preference and indifference relations are denoted P_i and I_i, respectively. We restrict attention to preferences that are continuous (i.e., for all $x \in X$, the sets $\{y \in X \mid y \; R_i \; x\}$ and $\{y \in X \mid x \; R_i \; y\}$ are closed), monotonic (i.e., for two bundles $x, y \in X$, if $x \geq y$, then $x \; R_i \; y$ and

[1] This canonical model has a long tradition in welfare economics. In the more recent literature, it is examined by Arrow (1963), Kolm (1972), Varian (1974), Moulin and Thomson (1988), and Moulin (1990, 1991), among many others.

if $x \gg y$, then $x\ P_i\ y$), and convex (i.e., for two bundles $x, y \in X$, if $x\ R_i\ y$ then $\lambda x + (1 - \lambda)\, y\ R_i\ y$ for all $\lambda \in [0, 1]$). Let $\mathcal{R}$ denote the set of such preferences.

An economy is denoted $E = (R_N, \Omega)$, where $R_N = (R_i)_{i \in N}$ is the profile of preferences for the whole population. Let $\mathcal{E}$ denote the class of all economies satisfying the preceding conditions. An *allocation* is a list of bundles $z_N = (z_i)_{i \in N} \in X^N$.[2] It is *feasible* for E if

$$\sum_{i \in N} z_i \leq \Omega.$$

We denote the set of feasible allocations for E by $Z(E)$. As we see in the following sections, each subfield of welfare economics addresses different questions in this model, and offers specific ways of solving them.

1.2 THEORY OF FAIR ALLOCATION

The theory of fair allocation (for a survey, see Thomson 2010), pioneered by Kolm (1968, 1972) and Varian (1974), looks for ways of allocating resources that are efficient, in the sense of Pareto, and fair. It turns out that, typically, fairness does not receive a unique interpretation in resource allocation models. As a consequence, the theory aims to identify all possible ways of capturing intuitions of fairness, define axioms that encapsulate these intuitions, and look at allocation rules that satisfy the axioms. An *allocation rule* is a correspondence S that associates to each economy E, in a domain $\mathcal{D} \subseteq \mathcal{E}$, a subset $S(E)$ of the feasible allocations.

The two allocation rules that have received the most attention are the egalitarian Walrasian and the egalitarian-equivalent rules. The first one, which we denote S^{EW}, selects all the allocations arising as competitive equilibrium allocations from an equal division of the social endowment of goods. The individual budget delineated by the endowment $\omega_i \in X$ and market prices $p \in \mathbb{R}^{\ell}_{+}$ is defined as

$$B(\omega_i, p) = \{z_i \in X \mid pz_i \leq p\omega_i\}.$$

Allocation rule 1.1 Egalitarian Walrasian (S^{EW})
For all $E = (R_N, \Omega) \in \mathcal{E}$,

$$S^{EW}(E) = \left\{ z_N \in Z(E) \mid \exists p \in \mathbb{R}^{\ell}_{+}, \forall i \in N,\ z_i \in \max|_{R_i} B\left(\frac{\Omega}{|N|}, p\right) \right\}.$$

A feasible allocation z_N for an economy $E \in \mathcal{E}$ is *(Pareto) efficient* if there is no feasible z'_N such that $z'_i\ R_i\ z_i$ for all $i \in N$ and $z'_i\ P_i\ z_i$ for some $i \in N$. Let $P(E)$ denote the set of efficient allocations for E. We can now define the second allocation rule. It selects all the Pareto efficient allocations having the

[2] The notation X^N is simpler than $X^{|N|}$ and is just as correct, as X^N is the set of mappings from N to X.

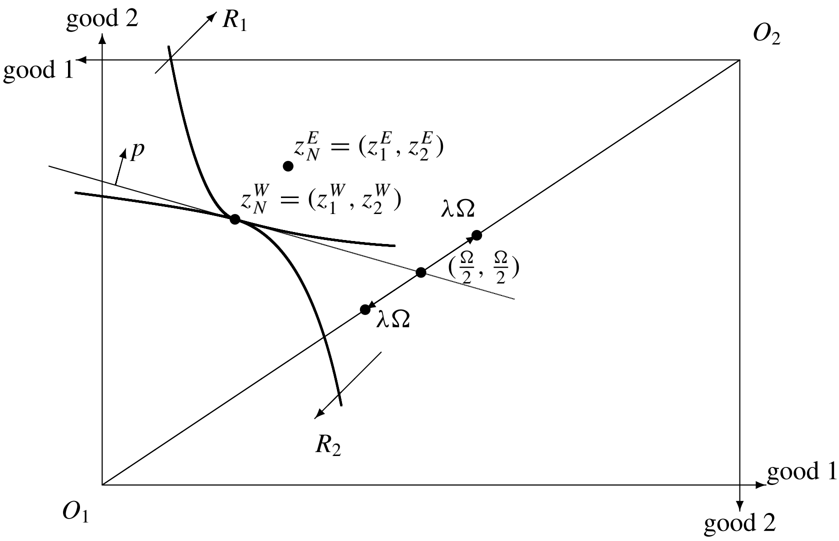

Figure 1.1. Illustration of the egalitarian Walrasian allocation rule

property that each agent is indifferent between his or her bundle and a fraction of the social endowment, the same for all agents.

Allocation rule 1.2 Egalitarian-Equivalent (S^{EE})
For all $E = (R_N, \Omega) \in \mathcal{E}$,

$$S^{EE}(E) = \{z_N \in P(E) \mid \exists \lambda \in \mathbb{R}_+, \forall i \in N,\ z_i\ I_i\ \lambda\Omega\}.$$

Figures 1.1 and 1.2 illustrate these allocation rules in the Edgeworth box. We have $N = \{1, 2\}$; preferences R_1, R_2 are represented by two indifference curves for each agent. The allocations represented in the figures are selected by the allocation rules defined earlier: $z_N^W = (z_1^W, z_2^W) \in S^{EW}(E)$ and $z_N^E = (z_1^E, z_2^E) \in S^{EE}(E)$.

Our theory has much in common with the theory of fair allocation, and we consider that we are mainly contributing to that theory. First, the (sometimes implicit) central ethical objective on which the theory is grounded is that of resource equality. In the simple model we use in this chapter, as well as in any other resource allocation model studied from that fairness point of view, if allocating goods equally were always possible and compatible with Pareto efficiency, then no other solution would be looked for. As explained in the introduction, we believe that equality of resources can be a key objective of a theory of fairness and social welfare, following the argument of Rawls' and other philosophers that social justice is a matter of allocating resources rather than subjective satisfaction or happiness.

Our work can be seen as an application of this approach in economic theory, to the limited extent that the models we study in this book are very partial descriptions of societies and that agents' preferences are a crude description

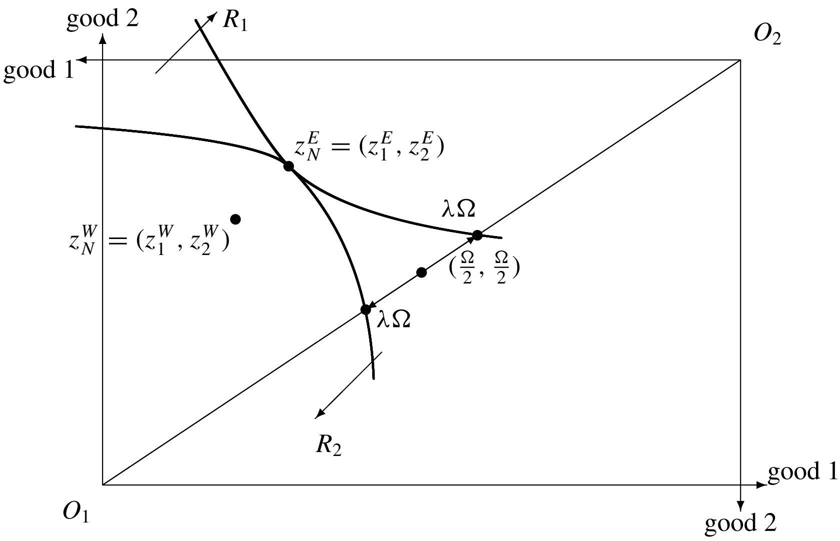

Figure 1.2. Illustration of the egalitarian-equivalent allocation rule

of their conceptions of a good life. On the other hand, our simple models are already sufficiently rich to prove that the seemingly simple notion of equality of resources can receive several interpretations, all of which are axiomatically justified.

The main difference between the theory of fair allocation in its current shape and our contribution is that a solution to a resource allocation problem in the former is an allocation rule, whereas in the latter it is a social ordering function. Let us explain this key point. When one studies allocation rules, the objective is limited to identifying the optimal allocations for each economy in a given domain, optimality being defined by a combination of efficiency and fairness axioms. In our approach, we look for *social ordering functions* (SOFs), which specify, for each economy, a complete ranking of the corresponding allocations.[3] Formally, a *social ordering* (for economy $E = (R_N, \Omega)$) is a complete ordering over the set X^N of allocations. A SOF $\mathbf{R}$ associates every economy E in some domain $\mathcal{D} \subseteq \mathcal{E}$ with a social ordering $\mathbf{R}(E)$. For $z_N, z'_N \in X^N$, we write $z_N \, \mathbf{R}(E) \, z'_N$ to denote that allocation z_N is at least as good as z'_N in E. The corresponding strict social preference and social indifference relations are denoted $\mathbf{P}(E)$ and $\mathbf{I}(E)$, respectively.

Let us illustrate this notion. The following example of a SOF also relies on the concept of egalitarian-equivalence and applies the leximin criterion to

[3] Tadenuma, Thomson, and other authors have studied complete allocation rankings based on fairness properties. They are precursors of the approach presented in this book. The rankings they obtain, however, all fail to satisfy the Pareto axioms on which we insist in the next chapters. See Chaudhuri (1986), Diamantaras and Thomson (1991), Tadenuma (2002, 2005), and Tadenuma and Thomson (1995).

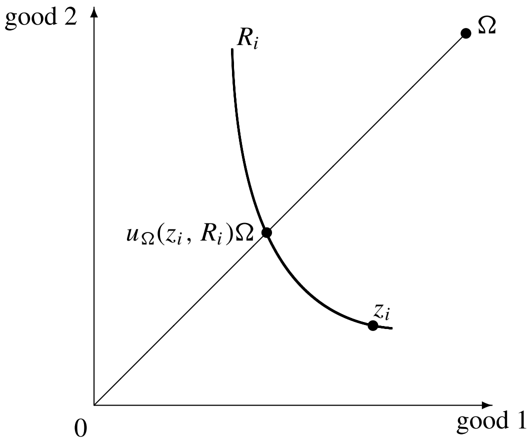

Figure 1.3. Computation of $u_\Omega(z_i, R_i)$

specific individual indices. The general definition of the leximin criterion is the following: for two vectors of real numbers a_N, a'_N, one says that a_N is better than a'_N for the leximin criterion, which will be denoted here by

$$a_N \geq_{lex} a'_N,$$

when the smallest component of a_N is not lower than the smallest component of a'_N, and if they are equal, the second smallest component is not lower, and so forth.

The specific individual indices to which the leximin criterion is applied by the SOF are defined as follows. They evaluate every agent's bundle by the fraction of Ω to which this agent is indifferent.[4] Indices of this sort are actually standard "utility" representations of individual preferences.[5] Formally, let us define the function $u_\Omega(z_i, R_i)$ by the condition

$$u_\Omega(z_i, R_i) = \lambda \Leftrightarrow z_i \; I_i \; \lambda\Omega.$$

We propose to call it the Ω-*equivalent* utility function, as it measures the proportional share of Ω that would give agent i the same satisfaction as with z_i. Figure 1.3 illustrates this notion.

When the leximin criterion is applied to Ω-equivalent utilities, one obtains the Ω-equivalent leximin SOF:

Social ordering function 1.1 Ω-Equivalent Leximin ($\mathbf{R}^{\Omega lex}$)
For all $E = (R_N, \Omega) \in \mathcal{E}$, $z_N, z'_N \in X^N$,

$$z_N \, \mathbf{R}^{\Omega lex}(E) \, z'_N \Leftrightarrow (u_\Omega(z_i, R_i))_{i \in N} \geq_{lex} \left(u_\Omega(z'_i, R_i)\right)_{i \in N}.$$

[4] This "egalitarian-equivalent" SOF was introduced by Pazner and Schmeidler (1978b), and, with an approach that is closer to ours, by Pazner (1979). The general idea of egalitarian-equivalence can be traced back at least to Kolm (1968), who attributes it to Lange (1936).
[5] See, e.g., Debreu (1959), Kannai (1970).

Of course, an allocation rule can be technically assimilated to a simple SOF for which the optimal allocations are socially strictly preferred to the nonoptimal ones, and for which all allocations are socially indifferent in each of these two classes. Such a SOF is obviously too coarse for many applications in which the fully optimal allocations cannot be obtained. Moreover, it violates the weak Pareto requirement, according to which an allocation is strictly better than another one as soon as each agent strictly prefers a bundle in the former allocation to the one in the latter. The SOFs we study in this book, such as the Ω-equivalent leximin, all satisfy this requirement.

Axiom 1.1 Weak Pareto
For all $E = (R_N, \Omega) \in \mathcal{D}$, *and* $z_N, z'_N \in X^N$, *if* $z_i \; P_i \; z'_i$ *for all* $i \in N$, *then* $z_N \; \mathbf{P}(E) \; z'_N$.

We think that SOFs may play a crucial role in the study of social fairness, after we take the implementation issue into account. Indeed, it is often the case that the set of allocations among which the policymaker must choose is so narrow that it does not contain any of the optimal allocations. Such constraints may come from asymmetries of information and, hence, incentive considerations (associated to the revelation of preferences). They may also be associated with the very nature of the problem. This is the case, for instance, if a status quo exists and the solution needs to be looked for in a neighborhood of it, or if the tools that the policymaker can resort to are limited – for instance, to linear taxation. Having a complete ranking of the allocations and maximizing it always leads to a well-defined solution (provided the set of implementable allocations is compact and the social ordering is continuous[6]) no matter which constraints turn out to be the binding ones.

Let us assume, for instance, that the social choice must be made among the set of allocations containing equal division and the allocations that can be obtained from it through trade at fixed price $\overline{p}$. Figure 1.4 illustrates, in a two-agent economy, the allocation $(\overline{z}_1^E, \overline{z}_2^E)$ that maximizes $\mathbf{R}^{\Omega\text{lex}}$ (both agents are indifferent between the bundle $\overline{z}_i^E$ they get and $\overline{\lambda}\Omega$, whereas all other allocations in the $\overline{p}$ trade line assign bundles that at least one agent finds worse than $\overline{\lambda}\Omega$). Observe that $(\overline{z}_1^E, \overline{z}_2^E)$ is inefficient and does not correspond to the first-best allocation for $\mathbf{R}^{\Omega\text{lex}}$.

In the following chapters, we give examples of applications of SOF maximization processes in the framework of second-best theory (that is, when the only constraint is that preferences or skills are the private information of the agents).

One may argue, however, that building SOFs is too demanding a task, as an extension of the allocation rule approach can be sufficient. Indeed, if the

[6] Continuity of the social ordering is actually not necessary. It is enough if the ordering can be constructed as the lexicographic composition of continuous orderings. Many SOFs that appear here, indeed, aggregate indices of well-being in a leximin way, and they are, therefore, lexicographic compositions of continuous orderings.

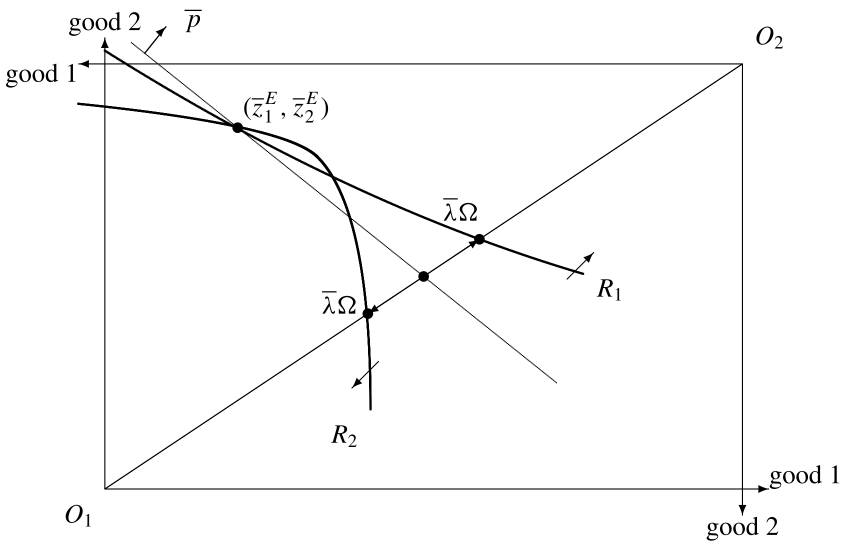

Figure 1.4. Best allocation for $\mathbf{R}^{\Omega\text{lex}}$ when only trade at fixed price $\overline{p}$ is possible

domain of problems for which an allocation rule is required to select optimal allocations is enlarged to include all problems that are likely to be faced by the policymaker (that is, all first-best and second-best problems, all possible reforms, and so on), then an allocation rule is as useful as a SOF.

Even if the preceding argument is perfectly valid, the social ordering approach remains justified, for the following reasons. First, as is well known from decision theory, as soon as the domain of problems is sufficiently rich, defining a consistent allocation rule is equivalent to defining a SOF. Second, from a technical point of view, studying an allocation rule in such a rich domain of problems requires identifying the set of feasible allocations in a sufficiently precise way, which may be infeasible (as is often the case when one looks at incentive constraints; an exception is Maniquet and Sprumont 2010). Focusing our attention on SOFs, therefore, guarantees the consistency of the policy recommendations that can arise from social welfare maximization under several kinds of constraints and simplifies our task.

Let us complete this section by pointing out another key aspect of the theory of fair allocation that is retained here – namely, its informational basis. Consider, for instance, $\mathbf{R}^{\Omega\text{lex}}$, which applies the leximin criterion to Ω-equivalent utilities. These "utilities," however, are mere indices representing ordinal and noncomparable preferences R_i. To make this clear, imagine that instead of starting our analysis at the level of preferences, we had introduced exogenous utility functions (measuring subjective satisfaction, for instance) $u_i : X \to \mathbb{R}$, and defined the set $\mathcal{U}$ of all utility functions representing continuous, monotonic, and convex preferences. In this alternative framework, an economy would be a list $E = (u_N, \Omega) \in \mathcal{U}^N \times \mathbb{R}^{\ell}_{++}$. In such a setting, our approach would simply ignore the properties of such utility functions other than the underlying preferences – that is, it would seek SOFs that satisfy the following axiom.

Axiom 1.2 Ordinalism and Noncomparability
For all $E = (u_N, \Omega)$, $E' = (u'_N, \Omega) \in \mathcal{D}$, *and* $z_N, z'_N \in X^N$, *if for all* $i \in N$ *there exists an increasing function* $g_i : \mathbb{R} \to \mathbb{R}$ *such that for all* $x \in X$, $u'_i(x) = g_i(u_i(x))$, *then* $z_N \, \mathbf{R}(E) \, z'_N \iff z_N \, \mathbf{R}(E') \, z'_N$.

Ordinalism follows from the fact that any utility function can be replaced by any strictly increasing transformation of it, so only the ranking of bundles matters. Noncomparability follows from the fact that those transformations of utility functions can differ among agents (we may have as many g_i functions as agents), so no common meaning can be attributed to utility levels or utility differences, in particular. Under ordinalism and noncomparability we do not lose any generality by defining economies directly in terms of preferences R_i.

At this point, we want to discuss an important clarification. Broadly speaking, the notions of well-being used by economists belong to three families. In the first family, well-being is measured in a way that does not take the individuals' subjective point of view into account. Examples include wealth and life expectancy. We can, of course, assume that doing better in these dimensions is better for the individuals, but how they trade off these dimensions against other components of well-being is not part of the picture. In the second family, all the attention is on subjective feelings or judgments. How agents trade off different dimensions of well-being is now taken into account, to the extent that nothing else enters the picture. Examples include happiness, subjective satisfaction, or utility. Our approach belongs to a third family. Individual preferences on how to trade off dimensions are respected, but we build indices of well-being in terms of quantities of resources that agents use to reach a given level of satisfaction. We return to this issue in Section 1.4.

1.3 ARROVIAN SOCIAL CHOICE THEORY

Contrary to the theory of fair allocation, Arrovian social choice looks for fine-grained rankings of the allocations. As explained earlier, our approach displays this feature as well.

The main difference between Arrovian social choice and our approach comes from the axioms we impose. The key axiom in the Arrovian tradition is the following independence requirement.[7] A SOF **R** satisfies this requirement if and only if the ranking between two allocations depends only on the individual preferences about these two allocations.

[7] It was called "independence of irrelevant alternatives" by Arrow (1963), a normative name that we do not retain, as it is controversial whether the concerned alternatives are indeed irrelevant.