

**INSTITUTE OF MATHEMATICAL ECONOMICS**

WORKING PAPERS

No. 333

**ZiF Research Group**

on

***PROCEDURAL APPROACHES TO CONFLICT RESOLUTION***

October 2001 - August 2002

Fascicle 5

**Social Choice and Just Institutions:  
New Perspectives**

by

Marc Fleurbaey

January 2002  
Revised March 2002



**University of Bielefeld**

**33501 Bielefeld, Germany**

# Social Choice and Just Institutions:

## New Perspectives

Marc Fleurbaey

University of Pau

marc.fleurbaey@univ-pau.fr

Jan. 2002

Revised, March 2002

**Abstract:** It has become accepted that social choice is impossible in absence of interpersonal comparisons of well-being. This view is challenged here. Arrow obtained an impossibility theorem only by making unreasonable demands on social choice functions. With reasonable requirements, one can get very attractive possibilities and derive social preferences on the basis of non-comparable individual preferences. This new approach makes it possible to design optimal second-best institutions inspired by principles of fairness, while traditionally the analysis of optimal second-best institutions was thought to require interpersonal comparisons of well-being. In particular, this new approach turns out to be especially suitable for the application of recent philosophical theories of justice formulated in terms of fairness, such as equality of resources.

## *Introduction*

Three main points are made in this paper. First, the aggregation of individual non-comparable preferences into reasonable social preferences is largely possible, in spite of Arrow's impossibility theorem of social choice. Second, this new approach to preference aggregation offers new possibilities for the analysis of the design of optimal institutions. Third, Rawlsian theories of justice will find here a valuable tool for concrete applications of fairness principles in the design of just institutions.

In short, the approach described here opens up new possibilities that, so far, have been largely ignored. The common view is that the aggregation of individual preferences requires interpersonal comparisons of well-being based on information that is not contained in non-comparable preferences. I will show that this view is not correct, and explain that it is due to unreasonable demands made on the aggregation procedure. Once the unreasonable requirements are relaxed, quite attractive aggregation possibilities appear.

Another common view, at least in economics, is that the design of redistributive institutions such as tax schemes and social security systems is best done by maximizing a social welfare function of the traditional kind, embodying interpersonal comparisons of well-being. This view is vindicated by the pessimistic interpretation of social choice that has just been described. Now, if interpersonal comparisons of well-being are no longer needed in social choice, they can be avoided in institution design, too, and this is why new possibilities are automatically created for this field as well.

A striking convergence between this new approach and Rawlsian theories of justice is that, in both, reference to interpersonally comparable subjective satisfaction is avoided in the discussion of the fair allocation of resources. In Rawlsian theories of justice,<sup>1</sup> this is based on the thesis that individuals should assume responsibility for their ends and ambitions in life, so that a low level of satisfaction in itself does not give any claim on social help.<sup>2</sup> In the traditional approach to social choice and welfare economics, the reluctance to

handle utility functions, as opposed to ordinal non-comparable preferences, mainly comes from the more pragmatic observation that no satisfactory method of measuring utility or satisfaction in an interpersonally comparable way has yet been developed. Be that as it may, this convergence should be good news for philosophers, who are usually at pains to imagine how their abstract principles can be applied to the design of just institutions. What is proposed in this paper is a closer cooperation between political philosophers, social choice theorists, and public economists, in the "assembly line" of just institutions. The first would provide the fairness principles. The second would formulate axioms embodying the principles, and derive social preferences. The third would find the institutions that are the best according to such social preferences.

### ***Hold it! Social choice is possible***

Social choice has been dubbed "the science of the impossible".<sup>3</sup> The root of this is obviously to be sought in Arrow's impossibility theorem.<sup>4</sup> Although the theorem was not taken very seriously by the specialists of welfare economics at the time of its publication, its influence has been growing because no good aggregation procedure has ever emerged from welfare economics, while the impossibility theorem has turned out to be very robust, and its main content, reproducible in every imaginable model. Another reason for this influence is that Arrow's seminal contribution, although negative in its substance, was also quite positive in terms of method. The concept of a *social choice function*, mapping various profiles of population preferences into "aggregate" social preference relations, was a beautifully elegant theoretical object, and the idea of relying on the axiomatic method in order to look for a good function was perfectly appropriate. The same remarkable contribution created the theory of social choice... and put it at a deadlock.

One may wonder how welfare economics would have evolved if the impossibility theorem had not been so influential. History cannot be replayed, but its mistakes can be corrected. It is indeed a mistake to believe that Arrow's theorem bars the way to reasonable

aggregation procedures. I will not explore here the historical reasons why most people have come to believe that, and I will focus instead on the substance of the matter.

Let the theorem be re-examined. The set up is the following. There is a set of options (social states, policies), and a finite population of individuals. Every individual has preferences over the set of options, so that, when considering any pair of options  $x, y$ , she can say whether she strictly prefers  $x$ , or  $y$ , or is indifferent between  $x$  and  $y$ . Formally speaking, by "preference" or "preference relation", I mean a binary relation "at least as good as", which is a complete preorder (that is, reflexive, transitive and complete).<sup>5</sup> And one says that  $x$  is strictly preferred to  $y$  when  $x$  is at least as good as  $y$ , whereas  $y$  is not at least as good as  $x$ ; and indifference between  $x$  and  $y$  means that any of these two options is at least as good as the other one. A *profile* of population preferences is a list describing every individual's preferences.

In this set up, the social choice problem is to define a "social" preference relation on the same set of options, as a function of the population's preferences. A *social choice function* is a mapping which defines a social preference relation for every profile of the population's preferences in some domain. Intuitively, the purpose of the social choice function is to make a synthesis of the population preferences. A social preference relation is, formally, the same kind of object as an individual preference relation, namely, a complete preorder over the set of options. But unlike individual preferences which are real characteristics of existing individuals, social preferences are a pure construct, and are meant to provide guidelines for social decisions. They should not be viewed as subjective preferences of some kind of collective entity.

In late formulations of Arrow's theorem, the following four conditions, or axioms, are imposed on the social choice function:

**Unrestricted domain:** The social choice function must be defined for every conceivable profile of population preferences.

**Weak Pareto:** When all individuals strictly prefer one option to another, social preferences must follow.

**Independence of Irrelevant Alternatives:** Social preferences on a pair of options should only depend on the population's preferences on these two options.

**Non Dictatorship:** No individual should impose her strict preferences on social preferences for all profiles of population preferences.

The theorem says that these four conditions are incompatible.

**Arrow's Impossibility Theorem:** *There exists no social choice function satisfying the four conditions, when the set of options contains at least three options, and the population has at least two individuals.*

I now proceed to show that the theorem is not so relevant as usually thought, because the set of conditions is unreasonably demanding. Of the four conditions, only the last one is totally unquestionable. No sensible aggregation procedure can ever let one individual always impose his strict preferences. In fact, one could readily require a stronger condition of *anonymity*, that is, require that social preferences do not depend on individuals' names, or, equivalently depend on individual preferences in a symmetrical way (everyone should have an equal "weight").

The other three axioms can be criticized, but to unequal degrees. Unrestricted Domain is unduly demanding in many contexts in which reasonable individual preferences cannot be anything. Economists are used, for instance, to work with individual preferences which satisfy particular properties (continuity, convexity, etc.) in addition to being complete preorders. Even in non-economic contexts, such as voting, preferences are usually shaped in special ways (a conservative voter does not rank a communist candidate above a social-democrat). This criticism is not very important, because Arrow's theorem is robust to many domain restrictions, in particular economic ones.

The Weak Pareto criterion can also be criticized, from non-welfarist quarters especially. Individual preferences may not be all that matters in terms of well-being. More basically, ordinary preferences cannot be taken at face value, because they may embody irrational impulses, framing effects, imperfect information, etc. All such criticisms, however, do not really call for abandoning the Pareto principle, but only for reformulating it in terms of an appropriate notion of well-being. As a consequence, I would like, from now on, to assume that individual preferences reflect individual well-being in an ethically acceptable way, whatever that is.

My main target of criticism will be the axiom of Independence of Irrelevant Alternatives (hereafter IIA). As Hansson wisely wrote, "Arrow's theorem is really a theorem about the independence condition".<sup>6</sup> I do not want to be understood as saying that IIA is an axiom based on unsound principles. It would certainly be a good thing if the social comparison of two options could depend only on individual preferences on these two options and on nothing else. That would make social choice a very simple matter. Very little information would be needed in the aggregation process. But simplicity and informational parsimony are not all that counts. Ethical relevance is also important. And IIA does wipe out ethically relevant information, in the social comparison of two options.

As an example of relevant information being wiped out, consider Ann and Bob, who, in the status quo, have the following bundles at their disposal. Ann has ten units of good 1 and two units of good 2, while Bob has three units of good 1 and eleven units of good 2. Would it be a good thing or not to transfer one unit of each good from Bob to Ann? This question can be embedded in a wider problem of social choice, concerning all conceivable distributions of the two goods to these two individuals. And we restrict our attention, as in IIA, to the formulation of social preferences over two options, the status quo and the transfer. Let us focus on the profiles of preferences in which both individuals are not satiable. They always prefer having more of each good. In all such profiles, Ann strictly prefers the transfer,

whereas Bob strictly prefers the status quo. But IIA requires social preferences to depend only on those pairwise preferences. Since the population preferences are invariable in the absence of satiation, social preferences should be constant too, for all profiles of non-satiable preferences. In other words, knowing that Ann prefers the transfer and Bob the status quo should be, on the basis of IIA, enough information. Then, when the population is not satiable, should social preferences favor the transfer, the status quo, or be indifferent?

With these scarce data, I am afraid indifference is the only reasonable option. Indeed, the status quo displays bundles (10,2) for Ann and (3,11) for Bob, and after transfer the bundles become (11,3) for Ann and (2,10) for Bob. There is a perfect symmetry between goods and people, which makes it impossible to prefer one option over the other.

But, and this my main point, there may be other relevant information that should influence social preferences. Suppose that in a particular profile P, both individuals have identical (non-satiable) preferences, and are indifferent between the bundle (10,2) and the bundle (3,11). Then, in this profile, the status quo provides Ann and Bob with bundles which both find equally valuable, whereas the transfer would make Bob unambiguously worse-off than Ann, since both would agree that his bundle would be less valuable, and he would envy Ann, in the sense of the envy test of the theory of fairness (an individual envies another if he would rather have the other's bundle). Preferences in profile P can even be assumed to be such that the status quo is a Pareto-efficient allocation,<sup>2</sup> whereas the transfer would destroy efficiency. Reasonable social preferences may certainly prefer the status quo on these grounds. But in another profile P', the situation might be reversed. Both might be indifferent between (11,3) and (2,10), which would justify the transfer, for similarly inspired social preferences.

In other words, it would be quite sensible for social preferences to rely on such information as "who prefers what bundle", or Pareto-efficiency of the allocation. But this is,



unjustifiably, forbidden by IIA. Notice that the kind of information that this example shows to be relevant belongs only to non-comparable ordinal preferences.

One may be worried that this example is about an economic problem, and wonder if IIA is not still valid for abstract, voting contexts. The economic issue of sharing resources is yet a central issue for the organization of societies, but let us turn to abstract contexts. It is true that in abstract issues less information about individual preferences is available anyway. For instance, it is impossible to know from individual preferences on abstract issues if one individual envies another's situation, a kind of information that is readily available in economic settings where options are concretely described in terms of who gets what. Then, when, in an abstract setting, little is known about options and preferences, is IIA acceptable, or is it still an excessively restrictive condition? The answer is likely to be the latter again. Consider two options,  $x$  and  $y$ . Suppose we know that Alice prefers  $x$  and Brian prefers  $y$ . According to IIA, this is enough information. What can social preferences be on such a poor basis? Probably indifference again. But suppose we now add the information that in  $x$  and  $y$  Alice is at her top and second best options, whereas Brian is at his worst and best options. Then it might become sensible to prefer  $y$ . Again, this additional information may be deemed relevant by reasonable social preferences, and it seems questionable to exclude it, as IIA does.

Now, the fact that IIA unduly excludes such information does not prove that such information would be sufficient to obtain a possibility instead of the impossibility theorem. The usual view is that bypassing the impossibility requires interpersonally comparable indices of well-being. I now proceed to show that this view is not acceptable.

Let us consider the abstract voting context first. This is the least favorable context, because, as explained above, preferences and options are described with minimal information. Nonetheless, there are several interesting social choice functions for this setting, which violate IIA on reasonable grounds, and satisfy other valuable properties. The best known is the Borda rule, which compares options on the basis of their average rank in people's

preferences. An option is better if it occupies, in average, a higher rank in individual preferences. This violates IIA because, quite reasonably, it examines whether going from  $x$  to  $y$  represents a big or small step in people's preferences; which depends on how many options are in between. A variant of the Borda rule, which seems to have been neglected in the literature, would compare options on the basis of their worst rank in people's preferences. An option is better, then, if it is farther from the worst option for each individual. This violates IIA because it compares options on the basis of how many options are still considered worse than them. The best variant of this rule considers the worst ranks lexicographically, in a "leximin" fashion. For instance, if  $x$  and  $y$  have fourteen individuals at their third worst option, but for  $x$  the next worst rank (for some fifteenth individual) is six whereas it is five for  $y$ , then  $x$  is to be preferred. Another important rule is the Kemeny-Young rule, which looks for a complete preorder which minimizes the total number of disagreements with the majority rule on pairwise comparisons. It has been axiomatically justified by Peyton Young,<sup>3</sup> and has recently received a vigorous philosophical defense by Mathias Risse.<sup>2</sup> This rule violates IIA because it discounts individual preferences which are outlandish, in the sense that they too often contradict the majority verdict in pairwise contests. It then often requires an extensive information about all majority comparisons before judging on two particular options.

Notice that, in the case when there are only two options in the set of options, the three rules coincide with the majority rule. In summary, these three examples are quite reasonable extensions of the majority rule, based on defensible violations of IIA. But the truth is that no revolution in welfare economics will ever come from these examples. A first reason for this is that the majority rule is a very poor device in distributive justice issues. Following the majority rule in questions of resource allocation is quite unwise because selfish majorities may too easily sacrifice minorities and deprive them of the little they have. In addition, majority cycles occur too often because any minority may have a majority raising against it, so that the extensions of the majority rules may too easily display a lot of indifference. This

kind of drawback is, however, mostly displayed by additive sorts of extensions of the majority rule, such as the Borda rule and the Kemeny-Young rule. The leximin variant of the Borda rule which has been introduced above does not behave so badly in distributive issues. For instance, if one has to share two indivisible units of good among two selfish individuals, the Borda rule and the Kemeny-Young rule are indifferent between all allocations, whereas, in contrast, the leximin variant of the Borda rule uniquely singles out the equal-split allocation as the best one.<sup>1</sup>

But even this more egalitarian rule is very poor and fails to give satisfactory solutions in many contexts, such as, for instance, the simple case when not all feasible allocations are considered in the set of options, or more standard economic problems with divisible goods. This kind of shortcoming is due to the impoverished information available in the framework. This is why economic issues have to be addressed with adapted tools, and cannot be subsumed under the political model of voting.

Let us now turn to an economic issue, such as, inspired by the above example of Ann and Bob, the problem of distributing divisible goods to a finite population of individuals. François Maniquet and I<sup>1</sup> have defended two particular kinds of social preferences for this particular case. The first one compares allocations by focussing on the worst share of the total available resources that an individual would accept to receive instead of her current bundle. For instance, if one allocation has one individual accepting to receive two percent of the total resources in exchange for his bundle at this allocation, whereas in another allocation no one would accept less than three percent, then the latter is to be preferred. Such social preferences bear a close relationship with the concept of fairness as "egalitarian-equivalence" proposed by Elisha Pazner and David Schmeidler,<sup>1</sup> because the best allocations among all feasible allocations

will be precisely the "egalitarian-equivalent" ones, in which everyone is indifferent between his current bundle and a given percentage (the same for all individuals) of total resources.

The second kind of social preferences is very similar, except that the shares of total resources are not considered for direct consumption, but for further exchange on the market, at some relevant prices. I will not go into the details of the definition here, but only mention that the best allocations for this second kind of social preferences are the competitive equilibria from equal-split, a famous concept of solution in the theory of fairness.

These two examples violate IIA, because they rely on much more information than pairwise preferences over two allocations. They depend on the shape of large parts of individual indifference curves. But, interestingly, it is enough for them to know the individual indifference curves at the contemplated allocations only, in order to be able to rank those allocations. This is a weakened form of IIA, which can be related to a weak independence axiom proposed, independently, by Hansson and by Pazner.<sup>1</sup> Is it reasonable to rely on indifference curves, instead of just pairwise comparisons? Yes, and this is precisely what the Ann-Bob example was meant to show. In that example, knowing the individuals' indifference curves at two allocations makes it possible to see whether in one allocation, for instance, both individuals have the same indifference curve, implying in particular that they are indifferent between their respective bundles, and whether, in the other allocation, one individual is envying the other. It also makes it possible to ascertain Pareto-efficiency of a particular allocation.

An exciting feature of this new kind of possibility is that a whole field of axiomatic research is opened. Because one should not be content simply with finding social preferences which satisfy all of Arrow's axioms except IIA. Arrow's axioms of Weak Pareto and Non-Dictatorship are quite minimal. In particular, the latter should be replaced by much more stringent equity requirements. As an example, consider the following requirements.

**Transfer Principle for Equals:** If two individuals have identical preferences and one has more of each good, it is socially as least as good to make a transfer of goods from him to the other, when ex post he still has more.

**Minimal Egalitarianism on Shares:** If all agents receive proportional shares of the total resources, it is socially at least as good to replace this by an equal split of this allocation.

These are just examples of equity conditions, among many others which may be imagined. In another paper,<sup>1</sup> I have shown that the Pazner-Schmeidler kind of social preferences defined above is singled out remarkably by the combination of these two axioms with Weak Pareto and the weak version of IIA described above. What is astonishing about this result is that such social preferences are of the maximin kind, since they focus on the worst-off individual, as measured by the share of total resources accepted, whereas none of the four axioms involved in this result does, by itself, imply such absolute priority for the worst-off. In particular, Transfer Principle for Equals is only about transfers in which what the recipient gets is equal to what the donor gives, and Minimal Egalitarianism on Shares is only about comparing perfect equality to inequality. None of these two axioms says anything about cases when the worst-off are helped at a great cost to the better-off.

This result is only an example and there exist other axiomatic justifications for this kind of social preferences and also for the other one.<sup>1</sup> The lesson which may be drawn here is that social choice is definitely possible. A mere weakening of IIA, and a quite reasonable one for that matter, is amply sufficient to give way to the construction of equitable and Paretian social preferences. I do not mean to ignore that ethics is plagued with dilemmas. It is certainly impossible to combine all desirable ethical requirements. But such impossibilities just circumscribe the domain of the possible, they do not nullify it. Even if we cannot fly, there

still are one thing or two that are worth doing on the ground. It is time for social choice and welfare economics to land.

### *Be fair with the fairness approach*

The title of Amartya Sen's beautiful Nobel lecture is "The possibility of social choice".<sup>1</sup> This author, along many others, has spent years defending social choice against what they called the "Arrow gloom". Their efforts have, however, concentrated on a different approach than the one highlighted above. They have accepted the thesis that social choice is impossible in absence of additional information about interpersonal comparisons of well-being, an information that is not contained in non-comparable individual preferences. But, biting the bullet, they defended the additional, reasonable thesis that interpersonal comparisons of well-being are not so hard to perform. It is not very difficult, in particular, to determine who should be given priority in societies with large inequalities.

Their second thesis is quite acceptable. But this does not impinge the fact that the first thesis is not acceptable. The examples of the previous section have shown that interesting social preferences can be based on ordinal non-comparable individual preferences. Moreover, the new approach that I have proposed on this basis may help defining social priorities in more convenient and more relevant ways. The relevance of the new approach will be discussed later. In this section, I would like to clarify a few theoretical points about the differences and similarities between the two approaches. For simplicity of reference, I will call Sen's approach the "comparability" approach, and the alternative one the "fairness" approach.<sup>1</sup>

An important similarity between the two approaches is that both reject IIA. But the essential difference is that they reject different *parts* of IIA. In order to understand this point, it is necessary to change the framework in order to have not only individual preferences but also interpersonally comparable indices of utility. Utility need not be understood here in the traditional welfarist sense, and can represent any concept of well-being that is judged relevant

for the evaluation of individual situations. Suppose, then, that we still have a set of options to rank, and that individuals are now described not by their preferences but by their utility measured on a numerical scale at the various options. This new setting can be viewed as containing more information, since from an individual's utility function, one can retrieve an individual preference relation by defining an option  $x$  to be at least as good as another option  $y$  for some individual whenever her utility at  $x$  is greater than or equal to her utility at  $y$ . In what follows individual "preferences" refer to this preference relation derived from the utility function.

In this new setting, one may still look for a social preference relation as a function of the population profile of utility functions. Arrow's axioms can still be formulated, with little change. The Unrestricted Domain now says that all individual utility functions are admitted. The other axioms do not even need to be rephrased. And Arrow's theorem is still valid.

Now, in this new formulation, it is important to notice that IIA says two things. First, it says that the comparison of two options should depend only on people's *preferences*, as opposed to the numerical values of utility. Second, it adds that the comparison must rely only on *pairwise* preferences about the two contemplated options. It is even possible to formulate this second restriction in a way that does not presuppose the first one: IIA requires the comparison to rely only on the numerical levels of utility at the two options. Let these two parts be formulated as separate axioms:

**Ordinal Non-Comparability:** Social preferences should remain the same when the profile of individual utility functions changes without altering individual preferences.

**Independence of Irrelevant Utilities:** Social preferences on two options should only depend on individual utilities at these two options.

Sen and others have very well analyzed this dual content of IIA. And they have argued that Ordinal Non-Comparability must be abandoned, whereas, in their approach, Independence of Irrelevant Utilities may be kept. This opens possibilities for social welfare

functions which compute social welfare for a particular alternative as a function of individual levels of utility at this precise alternative. In this way the second axiom is satisfied, since the comparison of two alternatives will depend only on the computation of social welfare at those two alternatives. Examples include the utilitarian social welfare function (focusing on the sum of utilities), the maximin (focusing on the minimum of utilities), etc.

The approach I present here does exactly the opposite. It retains Ordinal Non-Comparability, which implies that interpersonally comparable utility will be ignored, and non-comparable ordinal preferences will be exclusively considered. But it drops Independence of Irrelevant Utilities, and allows the social ranking of two options to depend on features of utilities (actually, preferences only) at other options. This difference may be illustrated with the Ann-Bob example. Recall that the status quo has Ann with bundle (10,2) and Bob with (3,11). Suppose that Ann's and Bob's utility levels at the status quo are respectively 4 and 5. And that the transfer of one unit of each good from Bob to Ann would reverse these figures. According to Independence of Irrelevant Utilities, this is enough information to make a social decision. Must we prefer the distribution of utility of 4 and 5 to the distribution with 5 and 4? We can see that both the utilitarian and the maximin social welfare function would indeed be able to adjudicate this case, and both would be indifferent between the two options. They are therefore unable to take account of the fact that, maybe, the status quo is an efficient allocation in which both agents are indifferent between their bundles, whereas the transfer would destroy efficiency and would make Bob have an unambiguously worst bundle. Such fairness considerations require more information about preferences than allowed by Independence of Irrelevant Utilities.

If one wants to take account of such fairness criteria, one has to drop this independence axiom. Once this is done, there is no logical necessity to drop Ordinal Non-Comparability any more, in order to obtain possibility results. As shown in the previous section, non-comparable preferences are then enough to get possibility results. It is only if one



assumes that Independence of Irrelevant Utilities is unquestionable that one can make the usual statement that social choice requires interpersonally comparable data on well-being. While this statement is often made, the implicit assumption on which it relies is commonly ignored. I hope that this analysis will have clarified this obscure point.

Another interesting fact, related to Independence of Irrelevant Utilities, obtains when one combines it with Pareto-Indifference, a condition saying that social preferences should be indifferent between two options when all individuals are indifferent, too. Under the Unrestricted Domain axiom, combining the two conditions (Pareto-Indifference and Independence of Irrelevant Utilities) yields the "welfarism lemma", due to Claude d'Aspremont and Louis Gevers,<sup>1</sup> and according to which options will be socially ranked on the basis of a unique ranking of vectors of utility levels, which will apply whatever the population profile. Let this ranking be called the "welfare ordering". The utilitarian or the maximin social welfare functions, for instance, always rely on the *sum* or the *minimum* of utilities, whatever the population profile. Comparing the sum or comparing the minimum define their respective welfare orderings.

It is interesting to note that such "welfarism" is built up in different ways by the two conditions involved. Independence of Irrelevant Utilities is the most important, which imposes that, for every pair of options, there exists a welfare ordering for that pair, independently of the profile. But the welfare ordering is specific to the pair, and it can be the utilitarian ordering for one pair, the maximin for another, etc.<sup>1</sup>

Pareto-Indifference merely adds that the welfare ordering must be the same for all pairs of options. In and of itself, Pareto-Indifference prevents social preferences to rely on non-utility features of options in the sense that, in a given profile, social preferences can be expressed by a welfare ordering. But since this welfare ordering is specific to the profile, this leaves a lot of possibilities on how to use utility information. The examples of social preferences presented in the previous section, for the distribution problem, do satisfy Pareto-Indifference, but use

utility information in a sophisticated way that makes it possible to consider quite elaborate fairness conditions.

When criticizing "welfarism", Sen<sup>2</sup> has targeted both Pareto-Indifference and Independence of Irrelevant Utilities, by advocating that non-utility features may be taken into account so as to violate both conditions. But, deep down, his "comparability" approach merely asks for a proper definition of utility. Once all relevant features of individual well-being have been put in the measure of utility, that is, for instance, when utility is redefined in terms of capabilities, then Sen is willing to accept both Pareto-Indifference and Independence of Irrelevant Utilities, leading to a revised, non-welfarist kind of "welfarism".

In contrast, by dropping Independence of Irrelevant Utilities, the "fairness" approach to social choice does no longer satisfy the welfarism lemma. This emphasizes the contrast with the "comparability" approach, and the fact that it does not merely apply a welfare ordering to exogenously defined interpersonally comparable utilities.

Nonetheless, it is clear that the fairness approach also performs comparisons between individuals.<sup>2</sup> For instance, the Pazner-Schmeidler example described above does compare the percentages of total resources that individuals would accept in exchange for their bundles. Moreover, it can be described as the application of the maximin criterion to such measures of individual "utility". Does not this blur the difference between the two approaches?

There are two essential differences between the ways the two approaches perform comparisons between individuals. The first one is about the general kind of object that is compared, reflecting potentially important philosophical divergences. In the "comparability" approach, comparison is about utility, whatever that is. In the "fairness" approach, the comparison is about bundles, as they are valued by individual preferences, with reference to equity principles. If, within the "comparability" approach it was decided that utility should merely reflect the value of bundles as assessed through individual preferences, then the two

approaches would, in their results, come close to each other. If, on the contrary, the "comparability" approach is about a more comprehensive notion of well-being, then the two approaches lead to social preferences that are radically different.

The second difference is about the method by which what is compared is defined. In the "comparability" approach, the job of defining utility is left to an outside field, such as moral philosophy. The analytical work performed by the theory of social choice is limited to the problem of aggregation, that is, of trading-off gains and losses in well-being for different individuals. In the fairness approach, the construction of what is being compared is justified *within* the theory, on the basis of axioms of fairness in particular. For instance, nothing a priori says that, in the distribution problem, the percentages of total resources that individuals would accept should play such a central role. It is only because of axioms such as Minimal Egalitarianism on Shares that they turn out to be relevant. In this approach, the axioms must, in one blow, define how to measure individual situations to be compared, and how to adjudicate conflicts of interests between individuals. In some cases it is actually possible to separate the role of different axioms into those which determine the measurement of individual situations and those which deal with inter-individual trade-offs. But the important point is that all this belongs to the same theory.

### *The best for the second-best*

After having clarified the similarities and differences between the more traditional "comparability" approach and the "fairness" approach, it remains to examine whether the latter deserves to be applied. In short, will it bring something new and valuable?

It must not come as a surprise that the answer put forth here will be positive, but it has two different parts. First, I will discuss the possibilities offered by this approach for the normative economic analysis of institutions, and show that the design of optimal institutions based on fairness principles is made much easier with this approach than with the previously existing methods. Second, I will examine the philosophical relevance of focusing on

individual bundles and fairness conditions, and relate this approach to the fairness principles which appeared in recent philosophical theories of justice. This two-part answer will not only establish a vague congruence between theories of justice and the ethical principles underlying the fairness approach, it will suggest that this approach provides a unique tool for the application of theories of justice to the design of just institutions.

Starting with institution design in general, one first has to recall that the difficulty of the exercise comes from the fact that individuals may be expected to react to the rules of the game by pursuing their own interest rather than a collective goal. Whether this makes achievable states of affairs radically unjust, or simply puts a constraint on the design of just institutions, has been the topic of a recent controversy raised by Gerald Cohen.<sup>2</sup> Here I am only interested in the search for *optimal* institutions in realistic conditions of human behavior, independently of whether such institutions deserve to be called "just" or not.

Economists have defined several kinds of incentive-compatibility constraints, depending on the amount of information about the population profile that is available when writing the rules of the game. I will focus on what seems to me the most relevant setting for the discussion of redistribution through devices such as taxes and benefits, social insurance and public assistance. This setting is usually called the "second-best" context. It consists in the following kind of assumption. The designers of redistributive mechanisms know the distribution of characteristics of the population, but do not know who has what characteristics in detail. They do not observe, at the individual level, personal characteristics such as education background, physical strength and other talents, and they also do not observe some of people's actions, such as their efforts at work, and also possibly the time they spend working. Nonetheless, the fact that they know the distribution of all these data over the population is important, because it makes it possible to compare the consequences of different institutions. For instance, they can forecast that if the tax scheme is such or such,

people with such or such characteristics will work and earn that much, yielding a global social state that can be evaluated if social preferences are well defined.

This is where social preferences of the various kinds discussed above are helpful. Since the social state of affairs induced by any precise institution can be accurately forecast, social preferences on social states may be used to derive social preferences on institutions.<sup>2</sup> Once social preferences on institutions are obtained, it is a pure optimization exercise to find the best institutions.

The second-best is opposed to the "first-best", which is the situation when the designers of the redistribution mechanisms know everything about everybody, so that they are only constrained in their control of the social state of affairs by the technical constraints of transformation and allocation of goods.<sup>2</sup>

How is institution design in the second-best context usually tackled? Most of the economic literature relies on the traditional utilitarian criterion. Even for a faithful utilitarian, the drawback of such applications is that the results about optimal institutions are quite sensitive to the particular choice of utility function used to measure individual well-being. For fixed individual preferences, one may indeed choose different utility functions representing those preferences, in particular with different degrees of concavity. Concavity of the utility function has important consequences on the degree of inequality aversion about income distribution, in utilitarian social preferences. Moreover, since it is already hard to choose a particular utility function, the exercise would be even more hopeless if individuals had different utility functions. As a consequence, most of the literature is confined to the unrealistic case when all individuals have the same preferences. And robust results are only those which do not depend on the particular choice of the uniform utility function of the population, which is very restrictive.

A more general "comparability" approach to the definition of social welfare can be resorted to, and again it has no other difficulty than the choice of a precise measure of

individual well-being. But this is no little difficulty. In addition, it does not lend itself easily to the incorporation of fairness considerations, as explained above in relation to the axiom of Independence of Irrelevant Utilities.<sup>2</sup>

If one is interested in applying fairness principles to the design of institutions, I claim that the "fairness" approach is the most convenient. This is not very surprising, since this approach is entirely oriented toward embodying fairness principles in full-fledged social preferences. But two other contenders have to be mentioned.

First, one may ask whether it is necessary to define social preferences over all allocations, as done in social choice in general. After all, one decision will be eventually made, leading to a precise social state of affairs (ignoring any intrinsic uncertainty). It would be enough to be able to select the best allocations among those which are achievable through feasible institutions. This kind of objection has repeatedly been made against social choice in general. Notice, nonetheless, that selecting a subset of allocations is still equivalent to defining a complete preorder over the set of allocations.<sup>2</sup>

But what is meant by this suggestion is that one could ignore fine-grained distinctions among bad allocations, such as required by the Weak Pareto axiom. This idea of directly looking for the second-best allocations is a priori a good one, but it seems hard to apply, for the following reason. Usually, fairness principles applied to allocations, as opposed to comparisons of allocations, are formulated in terms of some equality requested in some particular cases. For instance, individuals with identical preferences should end up on the same indifference curve. But the incentive-compatibility constraints that make up the second-best setting generally prevent such equalities to be achieved. The usual trick on which one falls back in such a situation is to replace the equality requirement by a priority statement, which takes the form of... a comparison between allocations. For instance, the transfer principle replaces a full equality request, as in Transfer Principle for Equals, and implies that some kind of equalizing

transfer produces a social improvement. Now, by falling back on comparisons of allocations, one is essentially turning to the social choice approach proposed here.<sup>2</sup>

Another contender is the method which consists in defining a pure fairness ranking, ignoring Pareto considerations, and then, applying this fairness ranking to the subset of second-best efficient allocations only, in order to avoid ending up with an inefficient allocation. This method is not unreasonable, and might sometimes yield interesting results.<sup>2</sup>

One may, however, wonder whether it is a radically different method, or a particular case of the method proposed here. Its main characteristic is that it entirely drops any Pareto requirement. It would still have to discuss social preferences on the basis of axioms, and this is not fundamentally different from the "fairness" approach in general, apart from this particular choice about Pareto, which should render the axiomatic analysis of social preferences more difficult rather than easier.<sup>2</sup>

To summarize, the main point of this section is that the social preferences constructed in the "fairness" approach to social choice can readily be applied to the problem of choosing the optimal institutions in second-best contexts. As explained in the next section, this should be very good news to philosophers who develop concepts of fairness.

### *Where philosophers go*

As explained above, the initial motivation for developing concepts of fairness has been quite different in economics and in philosophy. While economists merely fret about interpersonal comparisons of well-being and feel more at ease with bundles and preferences, philosophers such as John Rawls and Ronald Dworkin have come up with deep ethical justifications. They argue that social justice may be defined in terms of equality of resources provided to individuals, implying that individuals are left responsible for the way in which they transform resources into achievements and satisfaction. In the case when resources do

not benefit equally to differently gifted individuals, personal talents may be counted among the "extended" resources made available to individuals.

Such theories of justice are very close, in their formulation, to the fairness idea that individuals should receive bundles of equal value, as assessed through their preferences. In this section I want to show the potentialities and advantages of a rigorous method starting from basic principles, deriving social preferences from the basic principles, and ending with the analysis of optimal institutions for such social preferences.

Dworkin is the author who makes the most use of fairness concepts such as the equal-split auction or the envy test.<sup>3</sup> But he is obviously at pains when he has to describe what the best redistributive scheme should be. His favorite theoretical reference is a hypothetical insurance market on which individuals could buy insurance, including against personal handicaps (this involves a veil of ignorance in order to avoid that only affected people get insured, leading the redistributive agency to bankruptcy). But how does one relate the virtual outcomes of this hypothetical market to actual redistribution? The only answer that can be proffered is that the redistribution system should try to come as close as possible to the hypothetical ideal. Now, how does one define closeness? This must involve a notion of distance, allowing a comparison of allocations. We are then not very far from the social choice problem.

But Dworkin's reference to the insurance market is somehow half-way between a general formulation of principles and a concrete description of optimal institutions. Dworkin does also define general principles. First, he relies on the envy test in order to check equality of resources at some basic level. He then goes on and adopts the equal-split auction as an ideal mechanism which guarantees that the resulting allocation is envy-free. And he conceives the best application of the equal-split auction to the problem of personal talents as a formula of equal endowments on the hypothetical insurance market. In addition, he formulates two general principles in order to define equality of extended resources at a very



general level. First, the *final* allocation of resources should be 'endowment-insensitive', in order to avoid any penalization of people with bad personal characteristics. Second, it should be 'ambition-sensitive', in order to reflect people's choices based on their particular preferences.

Now, relying on the concept of envy and the last two principles, it is possible to show that Dworkin's insurance mechanism is a very poor device. This has been shown by John Roemer<sup>3</sup>, and I recently tried to analyze this problem in more detail.<sup>3</sup>

In a nutshell, this mechanism may lead to unacceptable transfers in the (frequent) cases of talents and handicaps which affect either people's productive abilities, or their marginal utility. In two other papers, François Maniquet and I have instead applied the two principles of endowment-insensitivity and ambition-sensitivity in a full-blown social choice analysis of the problem of redistribution when people have different labor-income preferences and unequal earning abilities. The results are interesting families of social preferences, which perform better than Dworkin's mechanism in terms of the above principles.<sup>3</sup>

As an illustration, consider the following axioms, applied to this context where individuals are described by their labor-income preferences and their earning ability. As in Dworkin's approach, it is assumed that earning ability is a characteristic which elicits social help when it is low. The first axiom reflects the endowment-insensitivity requirement quite clearly, by requiring to avoid any inequality between individuals who might differ only in their earning ability.

**Compensatory Transfer:** If two individuals have identical preferences and work the same amount, but have different consumptions, any transfer that reduces consumption inequality between them (leaving labor quantities unchanged) is socially acceptable.

To prevent any equivocation, it must be emphasized that such an axiom is not about institutions but about a pure social choice comparison of allocations. The way in which such allocations might be obtained is irrelevant at this stage. This axiom invokes a hypothetical

transfer, but at this stage one must not ask whether such a transfer is feasible or compatible with incentives. The question is only: If such a transfer *could* be done, *would* that be a good thing? The axiom answers that it would never be a bad thing (the axiom allows indifference and is therefore quite weak).

It is this first axiom which Dworkin's mechanism miserably fails to obey, because the hypothetical insurance market will lead people to maximize their expected utility over all possible earning abilities that they may have, and since an expected utility is a weighted sum of utilities, this will lead to utilitarian, not egalitarian kinds of redistribution.<sup>3</sup>

The second axiom reflects the ambition-sensitivity requirement, by forbidding any redistribution when all agents have the same earning ability. The *laissez-faire* allocation is then one in which unequal incomes reflect individual preferences about leisure and income.

**Laissez-Faire in Equal-Skill Economies:** If all individuals have the same earning ability, no redistribution is needed.

The last axiom is a mere separability requirement, whose purpose is to make sure that the social preferences will have a simple structure. It can also be justified on the grounds that individuals who are not affected by a change in the allocation of resources need not have a say in the evaluation of this change.

**Separability:** If an individual has the same labor-income bundle in two allocations, the social ranking of these two allocations should be the same if this individual did not exist.

A

remarkable fact<sup>3</sup>

is that these three axioms, combined with Weak Pareto and the Hansson-Pazner weak variant of IIA, entail that an allocation must be strictly preferred to another if its minimal "equivalent ability" is greater. The "equivalent ability" is a personal figure which measures, for any individual at any particular labor-income position, the earning ability that would provide the

same satisfaction as her current position to this individual, in absence of any tax and transfer (that is, if this individual had to live with this earning ability in a *laissez-faire* economy). As an example, consider an individual who works and pays taxes. Suppose he was proposed a job with a tax-free wage (and no restriction on his choice of labor time), and he would only be willing to accept it if he was proposed an hourly wage of at least  $w$ . Then  $w$  measures his "equivalent ability". Now, the *minimal* equivalent ability at a given allocation is simply the smallest value of individual equivalent abilities, among the population, at this allocation.

These social preferences, in other words, compare individual positions in terms of equivalent ability, and apply the maximin criterion to such measures of individual situations. Again, one obtains a strongly egalitarian kind of social preferences (the maximin) on the basis of limited redistributive principles. The axiom of Compensatory Transfer by itself does not imply any strong egalitarianism, and it is only in combination with the other axioms that it leads to the maximin.

Now, in the general method proposed here, the next step is to ask what kind of redistributive institutions would be optimal for such social preferences. Two striking facts can be derived.<sup>3</sup> First, if there are individuals with no productive ability, then the best redistribution scheme is the one which maximizes the minimum income guaranteed to all, and which may be distributed as a basic grant. But the second fact is that, if the minimal earning ability in the population (among actual, not equivalent, abilities) is greater than zero, then the best redistribution scheme does no longer maximize the basic grant. An interesting result, in this case, is that those individuals who have the minimal earning ability and work full time should not be taxed and should instead receive a subsidy, as in the Earning Tax Credit system. More strikingly, they should actually be given the *greatest* subsidy of the whole population. In other words, the optimal tax schedule is then very favorable to the hardworking poor.

Putting aside the substance of these results, the main methodological point I want to make here is that philosophers who think in terms of equality of resources have an important ally in the fairness approach to social choice. If only they could refrain their institutional imagination and remain in the field of general principles, they would find in social choice the means to develop concrete economic applications of their principles. The reason for this is that the fairness approach is perfectly adapted to the translation of their principles into axioms.

The "fairness" approach to social choice, insofar as it relies on non-comparable ordinal measures of individual well-being, is not appropriate for all theories of justice. For instance, it cannot accommodate theories of equality of opportunities in which individuals are not responsible for their "utility functions". But the focus of this paper on the use of non-comparable preferences should not hide the existence of a very general approach that encompasses all variants of social choice. And, although this seems not to have been done yet, there is no impossibility to develop, for instance, axiomatic studies of responsibility-sensitive social welfare functions based on comparable measurement of individual well-being.<sup>3</sup>

Moreover, the "fairness" approach is not only suitable for theories of equality of resources. Any theory which declares that individual levels of satisfaction are irrelevant to issues of justice can turn to this approach for the study of economic applications. In a previous paper, I tried to articulate a responsibility-sensitive theory of justice that would not involve delicate distinctions between what individuals control or do not control (such as theories of equality of opportunities), or between preferences and personal resources (such as theories of equality of resources). I argued, like Scanlon,<sup>3</sup> that social justice is only about a subset of individual achievements, excluding a private sphere in which unequal achievements are not problematic at the bar of justice. Individual responsibility can then be doubly effective

in the definition of justice principles, first by requiring to protect this private sphere, and second by making it socially important that individuals be given occasions to exercise choosing activities that shape their lives, without entailing inequalities in socially relevant achievements. It is quite plausible that, along these lines, levels of subjective satisfaction should belong to the private sphere. On the contrary, indifference curves at which the individuals end up are socially relevant achievements, if only in view of Pareto conditions, and social preferences based on the evaluation of lists of indifference curves (one for each individual) would be appropriate in economic applications of this theory. This is precisely what the "fairness" approach to social choice is able to do. The kind of social preferences described above, for the example of income redistribution, might be acceptable for this theory as well. The axiom of Compensatory Transfer is quite uncontroversial for egalitarians. The axiom of Laisser-Faire in Economies with Equal Skills is, however, more problematic because it would lead to accepting that idle individuals starve if their preferences induce them to do so. I would rather replace this axiom by one forbidding any discrimination between the "deserving poor" who do not work only because of their low productivity and those who do not work because, for some reason, they are averse to labor.

**No Undeserving Poor:** If two individuals have labor-averse preferences or a low earning ability, so that they refuse to work unless subsidized, and they do not work at the current allocation but have different consumptions, any transfer of consumption that reduces inequality between the two is socially acceptable

The difference between Compensatory Transfer and this new axiom is that the former applied to individuals with identical preferences, whereas the latter applies to individuals whose only common characteristic is that they refuse to work unless subsidized. The new axiom is not in outright opposition to Laisser-Faire in Equal-Skill Economies,<sup>3</sup> but goes in another direction, and forbids any discrimination among the poor based on their preferences about labor.

The replacement of Laisser-Faire in Equal-Skill Economies by No Undeserving Poor, in combination with Compensatory Transfer, Weak Pareto, Separability and the weak variant of IIA leads to social preferences which measure individual positions not by the equivalent ability, but by the "equivalent income".<sup>4</sup> That is the income that an individual would accept to have, instead of his current position, if he could stop working, or, rather, if his income did no longer depend on the amount of labor performed.

Such social preferences, by applying the maximin criterion to equivalent incomes, lead to favoring the redistribution scheme that maximizes the basic grant in all circumstances.<sup>4</sup>

The surprising feature of these results is that an axiom such as No Undeserving Poor, which merely requests equality *among* the poor, leads to advocating a maximal redistribution from the rich to the poor.

The conclusion that the basic grant should be maximized is surprisingly close to van Parijs's theory,<sup>4</sup> although the reasoning leading to it is quite different. van Parijs does not construct social preferences, and, in a breathtaking shortcut between basic principles and institutions, essentially argues that the basic grant is a measure of individual real freedom. This argument cannot convince those who think that leisure is a valuable resource. Now, measuring individual situations not by the basic grant, but by the "equivalent income", as suggested above, does take care of this consideration. In general, increasing an individual's leisure, for a given level of income, does increase her equivalent income (she would require a higher income to stop working). The fact that this leisure-sensitive measure of individual situations does nevertheless lead to maximizing the basic grant is all the more striking.

### ***Conclusion***

I will try to summarize the main points. First, there exists a new approach to social choice, which avoids resorting to interpersonally comparable measurements of well-being, and obtains possibility results by relaxing Arrow's axiom of IIA. This "fairness" approach

does not only provide possibilities, but allows the analyst to impose substantial equity principles on social preferences.

Second, such social preferences can immediately be used in the selection of second-best institutions, under incentive-compatibility constraints.

Third, and most importantly, the combination of the first two points provides new tools for the application of theories of justice in which it is considered irrelevant to examine individual subjective satisfaction.

This why, as announced in the introduction, I call for a closer cooperation between political philosophers, social choice theorists, and public economists. The fairness approach is a new bridge between philosophy and economics. It provides a rigorous method for the examination of the institutional implications of principles of justice that have so far been thought (in economics) to be inapplicable in second-best analysis.

In conclusion, I would also like to emphasize the importance of the first point for the general outlook of social choice. Even though many authors have already spent their energy fighting the "Arrow gloom", the fact that most of them accepted Arrow's own conclusion that interpersonally comparable measures of well-being were needed, substantially weakened their cause. Now, we can really say "social choice is possible", and hope that Arrow's theorem will, at last, be put back where it belongs, namely, among ordinary impossibility theorems that simply warn us that desirable, but often unreasonable, ethical requirements are sometimes mutually incompatible.

**Acknowledgments.** This paper is based on a part of a talk given at the conference "Procedural Approaches to Conflict Resolution", Bielefeld, Oct. 2001, and was presented at a workshop on equal opportunity, Bayreuth, Feb. 2002. I thank the participants at these two events, as well as participants at a philosophy seminar in the University of Montreal, for their reactions. I am grateful to François Maniquet for many discussions and a long collaboration on these issues.

- 
- 1 Such as John Rawls, *A Theory of Justice*, (Cambridge, Mass.: Harvard University Press, 1971); Ronald Dworkin, *Sovereign Virtue. The Theory and Practice of Equality*, (Cambridge, Mass.: Harvard University Press, 2000); Philippe van Parijs, *Real Freedom for All*, (Oxford: Oxford University Press, 1995).
  - 2 A low level of satisfaction must not be confused with a gloomy temper, which, as a personal characteristic equivalent to an internal (bad) resource, can elicit social help.
  - 3 See e.g. Peter J. Hammond, "Social Choice: the Science of the Impossible?," in G.R. Feiwel (ed.) *Arrow and the Foundations of the Theory of Economic Policy* (New York: New York University Press, 1987), pp.116 –134.
  - 4 Kenneth J. Arrow, *Social Choice and Individual Values* (New Haven:Yale University Press, 1951).
  - 5 The preference relation is reflexive if any option is as good as itself (a trivial property); transitive if whenever option x is at least as good as option y, and the latter is itself at least as good as option z, then x is at least as good as z; complete if for any pair of options, one of them must be at least as good as the other one.
  - 6 Bengt Hansson, "The Independence Condition in the Theory of Social Choice," *Theory and Decision* 4 (1973): 25-49, p. 25.
  - 7 A Pareto-efficient allocation with all individuals on the same indifference curve is, moreover, also a Walrasian equilibrium with equal budgets (also called sometimes an equal-split competitive auction).
  - 8 H. Peyton Young, Arthur Levenglick, "A Consistent Extension of Condorcet's Election Principle," *SIAM Journal of Applied Mathematics*, 35 (1978): 285-300.
  - 9 Mathias Risse, "Arrow's Theorem, Indeterminacy, and Multiplicity Reconsidered", *Ethics* 111 (2001): 706-734.
  - 10 There are three options in this example: 1) give two units to individual A; 2) give two units to individual B; 3) give one unit to each individual. Individual A ranks the three options in the following, decreasing, order: 1,3,2, whereas individual B ranks them as follows: 2,3,1. The average rank of any option



is 2, so that the Borda rule is indifferent between the three options. No option gets a strict majority over any other option, so that the majority rule yields indifference between the three options, and the Kemeny-Young rule follows. The leximin variant of the Borda rule focuses on the worst rank, and that equals 1 for options 1 and 2, but equals 2 for option 3, which is then strictly preferred.

- 11 Marc Fleurbaey, François Maniquet, "Utilitarianism versus Fairness in Welfare Economics," in M. Salles and J.A. Weymark (eds), *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls*, (Cambridge : Cambridge University Press, forthcoming); "Fair Social Orderings," mimeo (<http://aran.univ-pau.fr/ee/page3.html>).
- 12 Elisha Pazner, David Schmeidler, "Egalitarian Equivalent Allocations: A New Concept of Economic Equity," *Quarterly Journal of Economics* 92 (1978): 671-687.
- 13 Elisha Pazner, "Equity, Nonfeasible Alternatives and Social Choice: A Reconsideration of the Concept of Social Welfare", in J.J. Laffont (ed.), *Aggregation and Revelation of Preferences*, (Amsterdam: North-Holland, 1979).
- 14 Marc Fleurbaey, "The Pazner-Schmeidler Social Ordering: A Defense," mimeo.
- 15 Marc Fleurbaey, François Maniquet, "Fair Social Orderings"; Marc Fleurbaey, "The Pazner-Schmeidler Social Ordering: A Defense"; "Walras or Pazner-Schmeidler? Two Criteria for Social Decisions," mimeo.
- 16 Amartya K. Sen, "The Possibility of Social Choice," *American Economic Review* 89 (1999): 349-378.
- 17 The reader who is less interested in theoretical underpinnings than in consequences and applications may skip this section.
- 18 Claude d'Aspremont, Louis Gevers, "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies*, 44 (1977): 199-209.
- 19 To be rigorously precise, this holds true for every ordered pair, because the social ranking on well-being may not be neutral. For instance, it may say that  $x$  is at least as good as  $y$  if and only if the sum of well-beings is at least twice in  $x$ .
- 20 Notably in Amartya K. Sen, "Personal Utilities and Public

Judgements: or What is Wrong with Welfare Economics?,"*Economic Journal* 89 (1979): 537-58.

- 21 On interpersonal comparisons, see Marc Fleurbaey, Peter J. Hammond, "Interpersonally Comparable Utility", in S. Barbera, P.J. Hammond, C. Seidl, *Handbook of Utility Theory*, vol. 2, (Dordrecht: Kluwer, forthcoming).
- 22 See e.g. Gerald C. Cohen, *If You're an Egalitarian, How Come You're So Rich?*, (Cambridge, Mass.: Harvard University Press, 2000).
- 23 The astute reader will have noticed that this assumes that social preferences are anonymous. Since it is ignored who has what characteristics, social preferences may serve to evaluate the social state of affairs only if they focus on the distribution of individual situations, and do not want to favor Jones or Smith. As a counterexample, suppose that some particular redistributive scheme is favorable to hardworking people. If social preferences are not anonymous and want to favor Jones particularly, one cannot decide whether the redistributive scheme under consideration is good or not if one ignores whether Jones is hardworking or lazy.
- 24 This does not mean that this implies some authoritarian planning. As shown in famous theorems of welfare economics, the market may be used, supplemented by appropriate transfers.
- 25 In many cases, however, any (first-best or second-best) efficient allocation can be viewed as the maximum for a weighted utilitarian social welfare function, for some well chosen individual weights. In other words, for any efficient allocation one can find individual weights such that the contemplated allocation maximizes a weighted utilitarian social welfare function. Therefore, at a very general level, there is no restriction in considering only social preferences of the weighted utilitarian kind. But the weights are specific to the allocation. The difficulty is then to relate values of the individual weights to fairness principles, in order to get precise conclusions about the optimal institutions. A priori this seems an impossible task.
- 26 Selecting a subset of allocations is equivalent to defining the following social preferences: Any allocation that is selected is strictly preferred to any allocation that is not selected; all

selected allocations are equivalent; all non-selected allocations are equivalent.

- 27 In the first-best context, this difficulty is absent, and an impressive theory of first-best fair allocation has indeed developed. For a survey, see e.g. Hervé Moulin, William Thomson, "Axiomatic Analysis of Resource Allocation Problems," in K. J. Arrow, A.K. Sen, K. Suzumura (eds), *Social Choice Re-examined*, vol. 1, (London: Macmillan and New-York: St. Martin's Press, 1997). Unfortunately, the many positive results of this theory are useless for second-best applications.
- 28 An example of this approach is given by Yukihiro Nishimura, "Optimal Commodity Taxation for the Reduction of Envy", mimeo, who applies a fairness ranking due to A. Chauduri, "Some Implications of an Intensity Measure of Envy", *Social Choice and Welfare* 3 (1986): 255-270.
- 29 I am not aware of any axiomatic result along this vein, ignoring Pareto conditions.
- 30 See Ronald Dworkin, "What is Equality? Part 2: Equality of Resources", *Philosophy & Public Affairs* 10 (1981): 283-345; *Sovereign Virtue. The Theory and Practice of Equality*.
- 31 John E. Roemer, "Equality of Talent," *Economics and Philosophy* 1 (1985): 151-187.
- 32 "Equality among Responsible Individuals", in J.F. Laslier, M. Fleurbaey, N. Gravel, A. Trannoy (eds.), *Freedom in Economics*, (London: Routledge, 1998); "Equality of Resources Revisited", *Ethics*, forthcoming.
- 33 Marc Fleurbaey, François Maniquet, "Fair Social Orderings with Unequal Production Skills," *THEMA* 2000-17; "Fair Income Tax," mimeo, 2002.
- 34 For details, see Fleurbaey, "Equality of Resources Revisited".
- 35 See Theorem 1 in Fleurbaey and Maniquet, "Fair Income Tax". This result is just an example. With a different set of axioms, and in particular by replacing Laisser-Faire in Equal-Skill Economies by other expressions of ambition-sensitivity, one derives slightly different social preferences, which can also claim to represent equality of resources. See Fleurbaey and Maniquet, "Fair Social Orderings with Unequal Production Skills," for details, and Fleurbaey, "Equality of

Resources Revisited," for a non technical presentation of these alternative social preferences.

36 See Marc Fleurbaey and François Maniquet, "Fair Income Tax."

37 Responsibility-sensitive social welfare functions have been proposed by John E. Roemer, *Equality of opportunity*, (Cambridge, Mass.: Harvard University Press, 1998), and Dirk Van de gaer, *Equality of opportunity and investment in human capital*, Doct. dissertation, (K. U. Leuven, 1993). But they do not rely on axiomatic foundations, and can be criticized for their utilitarian flavor, which induces them to advocate utilitarian redistribution even when individuals have identical non-responsibility characteristics. For a comparison of alternative approaches, see Fleurbaey, "Equality among Responsible Individuals", and Walter Bossert, Marc Fleurbaey, Dirk Van de gaer, "Responsibility, Talent, and Compensation: A Second-Best Analysis", *Review of Economic Design* 4 (1999): 35-56.

38 Thomas S. Scanlon, "Preference and Urgency", *Journal of Philosophy* 72 (1975): 655-669.

39 In particular, there exist social choice functions which satisfy Laisser-Faire in Equal-Skill Economies, No Undeserving Poor, Weak Pareto, Separability, and the Hansson-Pazner variant of IIA.

40 See Theorem 3 in Fleurbaey and Maniquet, "Fair Income Tax."

41 Such social preferences may also claim to represent another approach to equality of resources. They belong to the particular family of social preferences presented in Fleurbaey, "Equality of Resources Revisited".

42 van Parijs, *Real Freedom for All*.