

# Economics and Philosophy

<http://journals.cambridge.org/EAP>

Additional services for ***Economics and Philosophy***:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



---

## A CONCEPT OF PROGRESS FOR NORMATIVE ECONOMICS

PHILIPPE MONGIN

Economics and Philosophy / Volume 22 / Issue 01 / March 2006, pp 19 - 54  
DOI: 10.1017/S0266267105000696, Published online: 18 April 2006

**Link to this article:** [http://journals.cambridge.org/abstract\\_S0266267105000696](http://journals.cambridge.org/abstract_S0266267105000696)

### How to cite this article:

PHILIPPE MONGIN (2006). A CONCEPT OF PROGRESS FOR NORMATIVE ECONOMICS. *Economics and Philosophy*, 22, pp 19-54 doi:10.1017/S0266267105000696

**Request Permissions :** [Click here](#)

# A CONCEPT OF PROGRESS FOR NORMATIVE ECONOMICS

PHILIPPE MONGIN\*

*Laboratoire d'économétrie, C.N.R.S. & Ecole Polytechnique, Paris*

---

The paper discusses the sense in which the changes undergone by normative economics in the twentieth century can be said to be progressive. A simple criterion is proposed to decide whether a sequence of normative theories is progressive. This criterion is put to use on the historical transition from the new welfare economics to social choice theory. The paper reconstructs this classic case, and eventually concludes that the latter theory was progressive compared with the former. It also briefly comments on the recent developments in normative economics and their connection with the previous two stages.

## 1. DIFFICULTIES SURROUNDING THE QUESTION, BUT WHY IT NEVERTHELESS DOES ARISE

In this paper I take up the challenge of discussing progress in normative economics. The difficulties surrounding the enterprise are obvious. First of all, it is notoriously hard to say what exactly normative economics is about – welfare or choice, value judgments or the study of value judgments, economic policy or armchair evaluation. Economic methodologists or theorists have provided grand statements on how normative economics should be separated from positive economics and applied economics; see

\* This paper supersedes an earlier one entitled “Is There Progress in Normative Economics?” (Mongin 2002). I thank the organizers of the Fourth ESHET Conference (Graz 2000) for the opportunity they gave me to lecture on this topic. Thanks are also due to J. Alexander, K. Arrow, A. Bird, R. Bradley, M. Dascal, W. Gaertner, N. Gravel, D. Hausman, B. Hill, C. Howson, N. McClennen, A. Trannoy, J. Weymark, J. Worrall, two anonymous referees of this journal, and especially the editor M. Fleurbaey, for helpful comments. The editor's suggestions contributed to determine the final orientation of the paper. The author is grateful to the LSE and the Lachmann Foundation for their support at the time when he was writing the initial version.

Keynes (1890), Robbins (1932), Samuelson (1947), Little (1950), Archibald (1959), to name but a few. However, these accounts are hardly compatible with each other, and it is not always clear how they relate to the work actually done in economics. The paper will adopt the following noncommittal view: the task of normative economics is to investigate methods and criteria for evaluating the relative desirability of economic states of affairs. This is a noncommittal statement because it does not say whether normative economics itself endorses the evaluations (and thus *makes* value judgments) or just explores the way of making them (and thus only *relates to* value judgments). Furthermore, it does not decide either whether a more desirable state is one involving more welfare, or more preference satisfaction, or more choice, or more of anything else. However, despite its utter generality, the definition is not vacuous. In particular, it makes it clear that normative economics has a teleological rather than a deontological structure, to use the familiar ethical distinction. That is to say, normative economics draws conclusions about the rightness of actions (here, policy arrangements) from a prior investigation of the desirability or “goodness” of economic states of affairs. The definition also encapsulates the claim that normative economics is primarily concerned with *evaluations*, and only secondarily with *recommendations* or *prescriptions*. It allows the economist to assess the functioning of markets without requiring that his evaluations be translated into specific policies. This is a view that I am going to take for granted here, although I realize that some might disagree with it.<sup>1</sup>

A second difficulty is that philosophers do not provide obvious guidance for the question I am tackling. They have nearly exclusively discussed progress in relation to science, while rarely contemplating the possibility that there is such a thing as *normative science*.<sup>2</sup> A further difficulty is that most of the available work on scientific progress deals with the empirical sciences; very little has been written on progress in logic and mathematics. Admittedly, even a suitable notion of *conceptual* progress for empirical sciences like physics or biology could prove valuable for my purpose. Unfortunately, philosophy of science does not have much to say about the more theoretical side of progress in the empirical sciences.<sup>3</sup>

Despite these bleak prospects, the question of progress in normative economics is a natural and even urgent one to investigate. The field exhibits

<sup>1</sup> More on the abstract issues of this paragraph in Mongin (2004).

<sup>2</sup> There is nonetheless a continental tradition of considering ethics as a normative science; see Kalinowski (1969) who traces it back to the Leipzig philosopher Wundt at the end of the nineteenth century. However, this tradition is hardly known outside France and Germany, and did not have much influence even there.

<sup>3</sup> This was emphasized by Laudan (1977: ch. 2), whose attempt to go beyond this negative diagnosis was meritorious but sketchy. Kitcher (1993: ch. 3) has further pursued the issue of conceptual progress in the empirical sciences.

a relatively simple pattern of development, and to the specialist at least, this pattern is both *intelligible* and *oriented*. Quite a few economists even believe that it is a *progressive* pattern – although they would find it uneasy to explain what they mean by that. I am interested in making sense of this intriguing view and assessing it. I offer this as an excuse for embarking on an adventurous paper.

## 2. THE HISTORICAL PATTERN OF NORMATIVE ECONOMICS

The historical pattern is easy to discern. The “economics of welfare,” as Pigou (1920) called it, reformulated and extended the patchy analyses of the social benefits of well-functioning markets that could be found in Marshall and other early neo-classicals. Pigou’s work is not only more focused than his predecessors’, but also much closer to the abstract definition of normative economics given above. Typically, it is clearer in distinguishing between the principles for evaluating economic states of affairs and the way these states of affairs come about in the market with or without state intervention (it is another contribution of Pigou that he identifies the corrective rôle of the state more precisely than his predecessors). However, when it comes to explicating his desirability concept, i.e., economic welfare, Pigou leaves the reader with insufficient guidance. In a related criticism, Arrow (1983: 18) noted that he had optimality conditions in mind but never properly explained what his maximand was. Whatever the exact meaning of his optimality conditions, he intended them to bear not only on the efficacy of the economy, but also on the distribution of income. Hence the easy and common reconstruction of Pigou’s *Economics of Welfare* as being utilitarian, a reconstruction which I believe requires further scrutiny.<sup>4</sup> This old-style welfare economics is the first form of normative economics. I will leave it aside for the rest of the paper.

The so-called *new welfare economics*, which crystallized in the 1930s and developed up to the 1950s, corresponds to the second historical form. It was much clearer than the older welfare economics about its premisses – prominent among which was what we now call the Pareto Principle<sup>5</sup> – and it eventually reached a conceptually clear separation between the optimality conditions themselves and their application to markets and economic policies. The main results obtained in these years were the fundamental theorems of welfare economics (I am using the modern terminology again for simplicity). The first fundamental theorem states that under mild conditions, a competitive equilibrium satisfies the

<sup>4</sup> Myint’s (1942) history of early welfare theories may be the last systematic account of *The Economics of Welfare*. The book cries out for a modern appraisal.

<sup>5</sup> “Individualism” in the older terminology of Bergson and Samuelson. Little (1950) is usually credited for the modern expression.

conditions for a Pareto optimum. The second fundamental theorem says that under more stringent conditions, any Pareto optimum can be obtained as a competitive equilibrium after the agents' initial endowments have been modified by suitable lump-sum transfers.<sup>6</sup> Using different conceptual and technical means, the new welfare economics was pursuing a slimmer version of Pigou's programme. Officially, it avoided the evaluation of income distribution, reserving it for the politician, the moralist, or the "economist qua citizen." The so-called Compensation Principle was an attempt to extend the optimality concept beyond the limits of the Pareto Principle while eschewing detailed distributive comparisons of the type exemplified by utilitarianism. The proponents of the principle believed that it was assertive enough to permit an evaluation of, say, the repeal of Corn Laws, although this measure had upset the income distribution between farmers, landowners, and wage-earners.

The third historical stage corresponds roughly to two different forms of normative economics, i.e., social choice theory on the one hand, and public economics on the other. It is often said that Arrow's *Social Choice and Individual Values* in 1951 struck a fatal blow to the new welfare economics. However, this claim cannot be interpreted as saying that social choice theory superseded welfare economics in its traditional rôle of assessing the working of markets and proposing improvements in terms of corrective taxes and the like. The objective of social choice theory set down by Arrow and further clarified by Sen's *Collective Choice and Social Welfare* (1970) is to investigate the various abstract methods of evaluating social states. Applications may or may not be market-related and enter the theory mostly by way of examples. From the 1970s onwards, it has been incumbent on the newly created discipline of public economics to discuss market optimality and policy corrections when the markets fail. Public economics has come to absorb most of the applied content of the "new welfare economics" that has survived criticism, so that there are currently two, quite distinct forms of normative economics being practiced in parallel. There may even be more than two if one takes into account inequality theory and poverty theory, which have developed in a relatively autonomous way for the last twenty years or so. Just by itself, this division process is enough to make the transition from the second to the third stage a complicated affair.

There is some evidence that normative economics might be undergoing another change. The bulk of social choice theory up to the mid-80s, and the whole of public economics roughly up to now, are *welfarist*. That is to say, they take the information provided by the individuals' utility functions to be necessary and sufficient data for the social evaluation

<sup>6</sup> Beginners sometimes believe that the two theorems taken together form an equivalence statement. This is not the case.

or the public decision.<sup>7</sup> This was the element of continuity between the third stage and the first two, as it were. From the point of view of social ethics, welfarism is a restrictive, and indeed conceptually problematic, principle to adopt. Internal criticism, especially in Sen's later work, as well as the recent dialogue between political philosophers and economists, have helped to bring this point home. Accordingly, some economists have started to reorient social choice theory in a non-welfarist direction. Sometimes they dispense altogether with utility functions, as they do when analyzing rights. More commonly, they supplement utility information with other sources, as when discussing talents and handicaps, opportunities and "capabilities." This theorizing is covered by fashionable labels such as "economic theories of justice" or "equity," which suggest a philosophical potential that welfare economics never claimed for itself, but there are also hints of implications and even applications, in the economist's specialized sense. So, arguably, normative economics is undergoing another metamorphosis. I hasten to add that not everybody in the field – even among those who contribute to reshaping it – would agree with the present suggestion. Some "equity" theories are still welfarist in the very sense of this paragraph,<sup>8</sup> and it is a fact that public economists are slow to catch up with the new developments. This said, nobody would deny that normative economics is on the move again and that welfarism is one of the major issues currently under discussion.

We may now be at the right historical distance to decide whether the third stage can be considered a progressive one. The present paper sets itself the more limited task of deciding whether social choice theory was progressive compared with the new welfare economics. Given the dissimilarities in scope I mentioned, the question can only relate to the *theoretical outlook* of the new welfare economics. A fuller assessment would have to include public economics, but I refrained from taking it into account here because of the complex preliminaries this would involve. While social choice theory emerged all of a sudden in Arrow's *Social Choice and Individual Values* in 1951, there is no pathbreaking work to signal the birth of public economics. It established itself as a field unobtrusively around 1970 by absorbing parts not only of welfare economics, but also of public finance, an ill-defined field which belonged more to positive than normative economics. At the time, both welfare economics and public finance had fallen into relative disrepute. Public economics combined whatever seemed worth taking in their legacy with scattered contributions such as Samuelson's analysis of public goods, Lipsey and Lancaster's work

<sup>7</sup> Here I follow Sen's usual definition of "welfarism," which goes in terms of utility functions. An alternative definition will be employed in section 7.

<sup>8</sup> Two examples are the "non-envy" and "egalitarian-equivalent" constructions; see Fleurbaey and Maniquet (1999) for a survey.

on “second-best” evaluation, Diamond and Mirrlees’s theory of optimal taxation. To make things even more complicated, public economics did not fully endorse the separation of welfare economics into old and new – while critical of both, it also borrowed something from each, and in particular sometimes revived a utilitarian style of evaluation. Although there exist valuable retrospectives,<sup>9</sup> I do not know of any authoritative summing up of these many connections. In contrast, there is a *received view*, which was established by Arrow and approved by his followers, of the connection between social choice theory and the new welfare economics. These writers claimed – and convinced the average economist to believe – that the new welfare economics was based on a hidden internal contradiction. Among other astonishing implications, Arrow’s theorem would lay bare the logical impossibility of a well-behaved Paretian social welfare function. The theorem would also point out the way of escape, which would consist in letting interpersonal comparisons of utility – be they utilitarian or of other kind – enter the social welfare function. This standard argument grounds the widespread idea that social choice theory *superseded* the new welfare economics.<sup>10</sup> This is an explicit claim of progress, which explains why I have centered the paper around it. Once it is clarified, I will compare it with the abstractly devised criterion of progress that is mooted in the next section. The major finding will be that the standard argument is ill-conceived but that the transition to social choice theory was progressive nonetheless, according to the criterion. It is as if the social choice theorists had seen the right move in the game, while giving for it a wrong reason.

### 3. A PROVISIONAL DEFINITION OF PROGRESS

I start by contrasting *intertheoretic* with *intratheoretic* progress. It is perhaps not too difficult to recognize advances made within the confines of a given theory when it is neatly structured – and this is the case of both social choice theory and the new welfare economics in its more abstract parts. There is a story of successive clarifications of the two fundamental welfare theorems, and a story of successive refinements of Arrow’s impossibility theorem. Both exemplify a form of progress in normative economics, but this is not the form I am interested in diagnosing, unless it interferes with the other form. Intertheoretic progress is what this paper is about.

When it comes to this kind of progress, controversy bursts out, and we can hardly do without an explicit definition. Making a bold attempt, I

<sup>9</sup> See in particular Hammond (1990) and Drèze (1995).

<sup>10</sup> Few works with the title “welfare economics” were published beyond the 1960s. The strongest ones, which are Feldman’s (1980) and Boadway and Bruce’s (1984), mostly consist of an admixture of social choice theory with public economics. The others, like de Graaff’s (1957) and Mishan’s (1969), or the later editions of Little (1950), are outdated restatements of pre-Arrovian welfare economics.

will say that a shift from a theory T to a theory T' is progressive if: (1) T' provides a solution to at least one unresolved problem of T; (2) T' provides a solution to the main problems that T had already addressed and resolved in its own way; (3) T' raises new problems and manages to solve at least one of them; (4) T does not satisfy the previous conditions with respect to T'.

This definition embodies the four ideas of (1) constructive criticism, (2) theoretical continuity, (3) independence, and (4) asymmetry, which are arguably the component parts of the common-sense notion of progress. Notice that if we take T and T' to refer to distinct variants of the same theory, we get a working definition of intratheoretic progress as a particular case. Importantly, the definition does not make particular reference to normative theories. The concept of problem-solving is broad – and vague – enough to apply to them as well as to theories in the empirical sciences and in mathematics. If one construes “problems” as either predictions to be confirmed or facts to be explained, one gets a definition similar to that of a progressive shift in Lakatos (1970).

Actually, something can be learned from the earlier debates surrounding Lakatos's methodology and Popper's (1963: ch. 10) related conception, which inspired it. This analogy suggests that there are two possibilities to consider for (1). Either the “unresolved problem” is already recognized by T and is very much like an *anomaly* accompanying T. Or it is not only solved but also pointed out by T', in which case it is like a *novel fact*. We might expect both kinds of situations to occur with normative theories. It is arguable that standard ethical rules, such as utilitarianism, are accompanied with anomalies.<sup>11</sup> In normative economics, the many difficulties surrounding the Compensation Principle were treated, at least initially, like anomalies. The case of Arrow's theorem, on which I will elaborate, illustrates the opposite model – that of a novel fact.

Something we learned from the discussions on research programmes is that it is most delicate to construe theoretical continuity appropriately. Instead of (2), I might have required that T' solve *all* the significant problems already solved by T. This would be asking too much, just as Popper's and Lakatos's famous requirement of non-decreasing content has proved to be too exacting. To say that just *one* of the earlier problems needs to be solved would be too lax. Accordingly, I remain vague in my clause (2) even if this is not very satisfactory. As for clause (3), it plays the same rôle as the requirement of added content in Popper and Lakatos, that is to say, it serves to exclude *ad hoc* modifications of T. Lakatos insisted that

<sup>11</sup> Consider for instance the discussion (and eventual dismissal) of fanaticism in Hare's (1976) utilitarian theory. The notion of anomaly is by no means limited to the empirical sciences. Mathematical theories can be accompanied with anomalies, as Lakatos's (1963–64) classic polyhedron example shows.



at least one of the independent predictions should be borne out by the facts, but Popper generally did not make this requirement.<sup>12</sup> My suggestion for (3) parallels Lakatos's condition, and is presumably open to the charge of disguised inductivism that was leveled against it by some Popperians.<sup>13</sup>

Here is where the analogy breaks down. The classic requirements of increasing testable content in Lakatos and Popper imply that there are *logical* relations between successive theories. On the simplest construal, T and T' will share a subset of their logical consequences. Once allowance is made for the fact that theories need auxiliary statements in order to deliver predictions, this straightforward conclusion need not hold anymore. But it is still the case that T and T' will be logically related, although in terms of other statements and in a possibly non-transparent way. Nothing of the sort is implied by the above definition; in fact, T and T' might respond to the same problems using entirely different means. For instance, it can happen that the problems that T was resolving actively are shown not to *arise* in T'. I would regard this as an instantiation of clause (2). Generally, when the notion of a successful prediction gives way to that of successful problem-solving, much – perhaps too much – flexibility is introduced. The theories in a sequence declared to be progressive according to (1), (2), and (3) may be related to each other in a number of ways. This is why I need (4) in order to include the commonsensical feature of asymmetry into my working definition of progress. The methodology of research programmes makes this clause redundant because of the logical relations already established by the analogues of (1), (2), and (3).

#### 4. THE SOCIAL-CHOICE-THEORETIC CRITIQUE OF WELFARE ECONOMICS: HISTORICAL LANDMARKS

##### 4.1 The general optimum and Bergson's welfare function

The new welfare economics isolated and placed considerable emphasis on the problem of determining the conditions for “the general optimum,” which it described as being a point of maximum social welfare. In essence, this was the problem of simultaneous maximizing the members of society's utility functions, given the interdependencies prevailing between producers and consumers and the constraints imposed on their available initial resources. The problem was resolved while assuming nothing about the cardinal measurability and interpersonal comparability of utility – that is, in contemporary language, by invoking only the Pareto Principle. For the present purposes, I will restrict attention to late restatements

<sup>12</sup> But see the requirement of empirical success in Popper (1963: 242–44).

<sup>13</sup> The issue of inductivism in the non-empirical sciences is touched on by Howson (1979), who also makes suggestions on how to apply the methodology of research programmes to non-empirical disciplines like mathematics.

of this solution by Bergson (1938), Samuelson, whose *Foundations of Economic Analysis* (1947) expands on Bergson's work, and Lange (1942), who takes a different approach. These three pieces exemplify the new welfare economics at its best and are thus suitable for a discussion of progress.

Bergson takes the step of discussing the general optimum conditions in terms "the Economic Welfare Function" (1938: 312), which takes as arguments the consumptions of commodities and expenses of factors (e.g., labour) of all the individuals. Symbolically,  $i = 1, \dots, n$  will denote the individuals,  $x_i$  the vectors of quantities consumed or expended by each  $i$ ,  $x = (x_1, \dots, x_n)$  the allocation vector of the economy, and  $E = E(x_1, \dots, x_n) = E(x)$  will represent Bergson's function. He makes the standard economic assumptions that  $E$  is increasing in individual consumptions and decreasing in individual expenses, and, at some point, that it satisfies the Pareto Principle, which he calls the Fundamental Value Propositions of Individual Preference (1938 : 318). Given the Pareto Indifference condition,  $E$  factors out in terms of the individual utility functions  $U_i$ , i.e., there exists another function  $W$  that is defined on vectors of utility values and satisfies the equation:  $E(x) = W(U_1(x), \dots, U_n(x))$  for all  $x$ . Adding the Strict Pareto condition, which makes the other half of the Pareto Principle, one concludes that  $W$  is increasing in each of its arguments. Bergson's contribution was to show that this thin set of assumptions was sufficient to obtain the already known conditions for the general optimum, i.e., that the marginal rates of substitution between commodities are equal from one individual to another, and similarly for the other relevant marginal substitution and transformation rates.

As Bergson also explains, more special conditions that appeared in the past can be traced back to supplementary assumptions imposed on  $W$ . For example, some of the marginal statements considered by "the Cambridge economists" – Pigou and his followers – depend on assuming the additive form  $U_1(x) + \dots + U_n(x)$ . For both the generic  $W$  and its specialized variants, Bergson derives marginal statements as the first-order conditions of a constrained maximization programme in which either  $W$  or its variants stands for the objective, and the technical possibilities set the constraints.<sup>14</sup> In the *Foundations* (1947: 229–53) Samuelson follows the same method of approaching the general optimum in terms of maximizing an objective function; hence the expression commonly used in the post-war years for the Paretian-inclusive  $W$ , "the Bergson–Samuelson social welfare function," At the early stage, neither author was clear about the extent to which  $W$  made interpersonal comparisons between the  $U_i$ . They knew that

<sup>14</sup> In keeping with the mathematical style of his time, Bergson used only intuitive arguments to conclude that his second-order conditions were satisfied, and relying as he did on the differential calculus, he had no way to handle corner solutions.

the Cambridge function did, since it was but a variant of utilitarianism, but it transpires from both the 1938 paper and the *Foundations* that they had not sorted out the case for the social welfare function *in general*.

This is an important claim for the discussion to come, and a possibly contentious one, so I will provide some textual evidence. Bergson remains cryptic throughout his paper about interpersonal comparisons of utility. He blurs the specific issue they raise by claiming that “value judgments” permeate all and every assumption underlying the Economic Welfare Function  $E$  (including the seemingly unproblematic Paretian conditions). The only place where he explicitly connects a “value proposition” with interpersonal comparisons is the passage on the Cambridge function (1938: 327). This obvious case does not help one to decide how he construes  $W$  more generally. However, once and almost inadvertently, he defines it in a way that *precludes* interpersonal comparisons of utility – he explains that the  $U_i$  can represent *indifference loci* (1938: 319). Samuelson is more informative than Bergson, and generally writes as if  $W$  did not make any interpersonal utility comparisons. In a passage I will return to later, he claims that

if we were to change from (the) set of cardinal indexes of individual utility  $U_1, \dots, U_n$  to another set  $U'_1, \dots, U'_n$ , we should simply change the form of the  $W$  function so as to leave all social decisions invariant. (1947: 228, notation adapted)

To paraphrase, when  $(U_1, \dots, U_n)$  is replaced by the cardinally different, but ordinally equivalent utility profile  $(U'_1, \dots, U'_n)$ ,  $W$ , will be changed into  $W'$  so as to leave the social preference unchanged. This is an exact rendering of Bergson’s claim that social welfare depends on indifference loci alone (and accordingly does not involve any interpersonal utility comparisons). However, Samuelson appears to retract this statement later, when he summarizes thus the case for social welfare functions:

Without a *well-defined*  $W$  function, i.e., *without assumptions concerning interpersonal comparisons of utility*, it is impossible to decide which of the [Pareto optima] is best. (1947 : 244, my emphasis)<sup>15</sup>

As in Fleurbaey and Mongin (2005), where this interpretation is presented in more detail, I conclude that Bergson’s and Samuelson’s early writings sorted out at most one of the two claims involved, i.e., that  $W$  did not logically *need* to make any interpersonal comparisons of utility. At that stage, the two economists had not decided whether or not  $W$  *should* normatively make such comparisons.

<sup>15</sup> My reading of this sentence hinges on “well-defined,” which suggests that interpersonal comparisons of utility are part of the definition of a “Bergson-Samuelson welfare function.”

Another landmark of the new welfare economics, Lange's (1942) paper has in common with Bergson's and Samuelson's work that it explores the logical possibilities of the Pareto Principle. Its second part contains a discussion of the general optimum that follows and actually improves on Bergson's, but the first part stands in sharp contrast with the latter's method of analysis. There, Lange introduced the (by now well-known) device of computing Pareto optima by maximizing one individual's utility function given that the technical possibilities are fixed *and* that the other individuals' utility functions are set at predetermined values. Thus, Lange also used the apparatus of constrained maximization, but differently from the other new welfare economists. The lasting importance of his method is that it does not require one to introduce a social welfare function in order to reach the marginal conditions for the general optimum.

#### 4.2 Arrow's theorem and the new welfare economics

Arrow's theorem has an immediate connection with Bergson's version of welfare economics, not with Lange's. It is no coincidence that the latter is mentioned only in passing in *Social Choice and Individual Values*, while the book makes the former the target of a lengthy and elaborate argument. Remarkably, after pointing out the wide generality of his notion of "social choice" in chapter I, Arrow chooses in chapter III to specialize it to welfare economics. This chapter introduces the conditions leading to the famous impossibility theorem not abstractly, but in terms of a "social welfare function," which he claims to share important features with Bergson's own function. The 1951 conditions are Universal Domain, Positive Association, Independence of Irrelevant Alternatives, Non-Imposition, Non-Dictatorship, and Arrow's definition of a social welfare function requires that this mapping deliver an ordering – I will call this implicit condition Social Ordering. For simplicity, I will use the following, slightly different set of five conditions: *Universal Domain, Weak Pareto, Independence of Irrelevant Alternatives, Non-Dictatorship, Social Ordering*. This list emerged from the 1963 revision and has since become standard. For a technical wording of each condition and proof that they are incompatible, the reader is referred to Sen's (1970) authoritative treatment.

The discussion of Bergson continues throughout Arrow's book, recurring in chapter IV on the Compensation Principle, and eventually culminating in chapter VI. At this juncture, Arrow goes beyond his initial claim that Bergson's function is *analogous* to one of his social welfare functions. He contends that it *is* in effect one of them, with the striking consequence that it falls prey to the impossibility theorem:

Mathematically, the Bergson social welfare function has... the same form as the social welfare function we have already discussed... Hence, the Possibility Theorem... is applicable here; we cannot construct a Bergson

social welfare function . . . that will satisfy Conditions 2–5 and that will lead to a true social ordering for every set of individual tastes. (1963: 72)

This is a crucial statement to understand the connections, both historical and logical, between the new welfare economics and social choice theory.

On a few occasions in *Social Choice and Individual Values*, Arrow goes even beyond the stage of rejecting Bergson's version of the new welfare economics. He suggests that his refutation makes the search for optimum conditions generally meaningless:

We may . . . doubt that any study of maximal alternatives will actually be useful in studying those aspects of social choice which are directly related to consumer's (and worker's) choice. (1963: 37)<sup>16</sup>

But there cannot be such a straightforward implication from Arrow's initial argument to this bold suggestion. I have stressed that Lange's derivation of the marginal conditions does not depend on using social welfare functions, which makes it immune to the attempted refutation. One interpretation of Arrow's quote is that he viewed the study of the general optimum as being *only a preliminary stage* in the construction of a social welfare function. In itself, this view would be hard to defend. Clearly, the marginal conditions have an interest by themselves, even if they do not inform us about the more difficult cases calling for distributional considerations. There is a further reason to doubt that Arrow seriously entertained the strong conclusion suggested by the quote – it would imply that the important work he did to improve on the two welfare theorems was pointless.<sup>17</sup> Having cleared up a possible misunderstanding, I return to the real object of Arrow's critique, which is the Bergson–Samuelson social welfare function.

## 5. THE SOCIAL-CHOICE-THEORETIC CRITIQUE OF WELFARE ECONOMICS: DEVELOPMENTS AND CONTROVERSIES

### 5.1 Arrow's argument against Bergson

Arrow's rejection of Bergsonian welfare economics depends on establishing that the Bergson–Samuelson function  $W$  is not only related to, but identical with, a social welfare function in *his* sense. This conclusion requires three steps, the first and the second of which appear to be unproblematic. The first step is purely semantic. Arrow's own function comes with a privileged interpretation of the individual preference relations it depends on – they are meant to represent the individuals' evaluations of social states, as influenced by their "values" (1963: 22).

<sup>16</sup> The same idea occurs in Arrow (1963: 63–64), where, however, it is significantly qualified.

<sup>17</sup> Arrow's major contributions to Paretian welfare theory took place roughly at the time of *Social Choice and Individual Values*. See his *Collected Papers*, vol. 2, especially ch. 2.

Bergson, and welfare economists generally, analyze social states in terms of individual consumptions and supplies of factors, and their notion of a utility function is meant to reflect the individual's ordinary, unelaborate preferences – his “tastes” as opposed to his “values” in Arrow's terminology (p. 23). As the book points out, this semantics can be accommodated by the social welfare function viewed as a purely formal object. Where an objection could arise, however, is with the Universal Domain condition. If “tastes” are construed according to standard microeconomics, i.e., as the individual's preferences varying positively with his consumption and negatively with his expenses, and depending on nothing else, a heavy restriction follows on the set of available preference profiles. Hence a second, purely logical step, which consists in showing that the impossibility theorem still holds despite the restriction (“Possibility Theorem for Individualistic Assumptions,” 1963: 63).<sup>18</sup> In the sequel I will refer to the new domain condition as *Modified Universal Domain*.

The ground is now cleared for the third and most problematic step, which is to defend the other conditions in terms of the general objective and privileged interpretations of Bergsonian welfare economics. Arrow (1963: 73) is disappointingly brief when it comes to this step. Essentially, he contents himself with reminding the reader of the general normative plausibility of the conditions – he had already defended them when introducing them formally. This appears to be an ineffective argumentative move. Given the task that Arrow had set for himself, he should have combined the logical use of his theorem with a specific *ad hominem* argument, to the effect that Bergson had implicitly accepted Non-Dictatorship and – above all – Independence of Irrelevant Alternatives. There are of course no questions with Social Ordering and Weak Pareto since they are contained in Bergson's statement of the *W* function.

Not surprisingly, the welfare economists plunged into the breach. Little (1952), Bergson (1954), and Samuelson (1967), conceded that the theorem was perhaps applicable to politics, although they would not feel entirely secure about this, but claimed most strongly that it fell outside their field. “We must conclude that Arrow's work has no relevance to the traditional theory of welfare economics, which culminates in the Bergson-Samuelson formulation,” said Little (1952: 141).<sup>19</sup> “I agree with Little in barring Arrow's theorem from welfare economics,” added Bergson (1954:

<sup>18</sup> This variant result justifies the earlier cryptic comment in the book that “the current analysis of maximal social states is applicable precisely when it cannot serve the function of a preliminary to a complete enumeration of the social ordering” (1963: 37).

<sup>19</sup> Baumol's early review of *Social Choice and Individual Values* had already set the pace: “This result is less disastrous for welfare theory than might first appear” (1952: 110).

247).<sup>20</sup> "I export Arrow from economics to politics because I do not believe that he has proved the impossibility of the traditional Bergson welfare function of economics," wrote Samuelson in the most famous paper of this series, "Arrow's Mathematical Politics" (1967: 42).<sup>21</sup> Later texts in welfare economics have often taken for granted the political interpretation of the impossibility theorem, as if it provided a satisfactory compromise between Arrow and his opponents. The usual approach goes as follows. The politically interpreted social welfare function decides *which* of the many Pareto optima should prevail; then, in accordance with the second welfare theorem, society entrusts the market with the task of implementing the selected optimum. In the end, the social choice of a Pareto optimum is constrained by Arrow's strictures, but this is due to an intervening electoral stage, and not to a possible failure of Paretian economics.<sup>22</sup> This approach concedes only *indirect* economic relevance to the impossibility theorem. It takes for granted the arguments promoted by Little, Bergson, and Samuelson to downplay the applicability of the theorem to social welfare functions as these economists conceived of them. I will review these arguments now.

## 5.2 The profile argument and the controversy of the 1970s

The first objection, which Little (1952) and Samuelson (1967) especially emphasized, was that the very notion of an Arrow function, as defined on a *set* of many preference profiles, made no sense in welfare economics; and similarly for the conditions put on this function that involve considering several profiles at a time. Indeed, Little and Samuelson argued that welfare economics was restricted to *given* individual tastes, which meant, in Arrow's framework, a *unique* preference profile. According to the argument, welfare economics comparisons bear only on changes in either the physical variables, such as individual consumptions, or the technological parameters, such as the firms' production possibilities. This can be recast mathematically as follows: the relevant social welfare function is *a composed function and not a functional*. The standard notation  $W(U_1, \dots, U_n)$  equivocates between the two senses because it could mean either:

$$(U_1, \dots, U_n) \rightarrow W(U_1, \dots, U_n)$$

or:

$$(x_1, \dots, x_n) \rightarrow (U_1(x_1), \dots, U_n(x_n)) \rightarrow W(U_1(x_1), \dots, U_n(x_n))$$

<sup>20</sup> Bergson's late restatements (1966 and 1976) uphold the same strong conclusion.

<sup>21</sup> Revisiting the Arrow-Bergson controversy, as well as his own controversy with social choice theorists, Samuelson (1981, 1987) came up with essentially the same claim.

<sup>22</sup> Feldman's (1980) text illustrates this double-sided approach very clearly.

It is the latter mapping which welfare economists have in mind, and they have no use for the former.

As it turned out from later discussions, the profile argument was not powerful enough to save welfare economics from Arrow's onslaught. To define a "social welfare function" on a set of many preference profiles would be immaterial if the conditions imposed on the function did not entail comparisons between several of these objects. Sen (1977, in 1982: 251–56) was the first to make this observation, which reduces the scope of the disagreement to the conditions themselves, and specifically to the subclass of those which are involved in the making of interprofile comparisons. The 1951 version had one too many of those problematic conditions – Positive Association, which gave way to the more familiar Weak Pareto in the 1963 version and ensuing texts.<sup>23</sup> What remains objectionable is the pair of conditions Universal Domain, in either its initial or adapted form, and Independence of Irrelevant Alternatives. The former provides the stock of profiles between which the latter allows one to make interprofile comparisons.<sup>24</sup> But crucially, the work done by social-choice theorists in the 1970s established that both conditions could be replaced by new ones stated for a *single* profile, leading to the reappearance of the impossibility theorem in this less controversial framework.<sup>25</sup> I will denote the single profile analogues of Modified Universal Domain, Independence of Irrelevant Alternatives, and Non-Dictatorship, by *Single Profile Modified Domain*, *Single Profile Neutrality*, and *Single Profile Non-Dictatorship*, respectively. Social Ordering and Weak Pareto do not need replacing because they are formulated identically for either one profile or many at a time. Around 1976–80 the three novel conditions displaced Arrow's initial ones as the focus of attention, and a fierce controversy took place between those social choice theorists who had promoted them and Samuelson, who acted as the only spokesman for the welfare economics camp.<sup>26</sup> Fleurbaey and Mongin (2005) have reappraised the controversy in full detail, and I will now report on some salient conclusions from this study.

The first variant condition, Single Profile Modified Domain, decomposes into the assumption of a rich domain of physical quantities and that of a given preference profile of individual preference that satisfies

<sup>23</sup> Thus, Little's (1952: 141) attacks on Positive Association proved to be ultimately vain.

<sup>24</sup> In contrast, Non-Dictatorship eschews the profile criticism. Dictatorship is defined across profiles, hence questionable, but just for that reason, *Non-Dictatorship* is relatively weak and acceptable.

<sup>25</sup> Kemp and Ng (1976) and Parks (1976) were the first to prove this result. Sen (1977), Pollak (1979), Roberts (1980) developed it further.

<sup>26</sup> It is surprising that Bergson and Little remained silent on such an important occasion. The welfare economists could also rely on the support of Mayston (e.g., 1982), but his work was unfortunately disregarded.



the standard economic assumptions. Its purpose was to create a common ground between the opposite camps. The third variant condition, Single Profile Non-Dictatorship, was more contentious. Commonsensically, dictatorship relative to a given profile is less unpalatable than it would be on a set of many profiles. However, after brief skirmishes around this issue,<sup>27</sup> the welfare economists conceded Single Profile Non-Dictatorship. The controversy focused almost exclusively on Single Profile Neutrality, whose technical rôle in the new framework corresponds to that assigned to Independence of Irrelevant Alternatives in the original one. This condition stipulates that if  $x$  and  $y$  are located on the individuals' preference maps exactly as are two other states  $w$  and  $z$ , then  $x$  and  $y$  may be replaced by  $w$  and  $z$  in the social preference, i.e., *society ranks the first pair exactly as it does the second*. Even for just one profile, this is a formidable assumption to make, as Samuelson was quick to point out. Take an "ethical observer" (Samuelson's personification of social preference) who must allocate 100 chocolates between two individuals:

What is the meaning of [Single Profile Neutrality] in this context? It says, "If it is ethically better to take something (say 1 chocolate or, alternatively, say 50 chocolates) from Person 1 who had all the chocolates in order to give to Person 2 who had none, then it must be ethically preferable to give all the chocolates to Person 2". One need not be a doctrinaire egalitarian to be speechless at this requirement. Is it "reasonable" to put on an ethical system such a straightjacket? Few will agree that it is. (1977: 83)

To connect Samuelson's example with the abstract condition, denote by  $x$  and  $y$  the allocation vectors  $(100, 0)$  and  $(99, 1)$ , where the components refer to numbers of chocolates consumed by 1 and 2, in that order. Society has the same preferences between  $x$  and  $y$  as between  $z = x = (100, 0)$  and  $w = (0, 100)$ , hence if it prefers  $(99, 1)$  to  $(100, 0)$ , it must also prefer  $(0, 100)$  to  $(100, 0)$ . Evidently, this conclusion defeats the egalitarian intent of the initial preference statement. No more than this little example is sufficient to deprive Single Profile Neutrality from its normative appeal as far as distributive issues are concerned, i.e., for welfare economics. Although this would have been possible, Samuelson did not adapt his counterexample to the political context. Such restraint is consistent with his long-standing view that Arrow's work is at least relevant to "mathematical politics."<sup>28</sup>

Persuasive as it is, Samuelson's example was not up the challenge posed by the single profile impossibility theorem, since the crucial question for the welfare economists was not to decide whether they *should* accept Single Profile Neutrality, but whether they *had* accepted it,

<sup>27</sup> See Little (1952: section 2) and Bergson (1954: 237).

<sup>28</sup> Turning Samuelson against himself, Pollak (1979) argued that if Single Profile Neutrality is objectionable in welfare economics, it may also be in relation to political or judicial rules.

possibly without noticing. In order to compare this condition with the welfare function  $W(U_1(x), \dots, U_n(x))$ , I will represent the given profile of preference relations in Single Profile Modified Domain by the set of all ordinal transforms  $(\varphi_1 \circ U_1, \dots, \varphi_n \circ U_n)$  of a given utility profile  $(U_1, \dots, U_n)$ , where the  $\varphi_i$  are increasing real functions and the  $U_i$  satisfy the relevant economic restrictions. Once this notational step is performed, it turns out that there are three possibilities for  $W(U_1(x), \dots, U_n(x))$ , each with distinctive consequences:

- (1)  $W(U_1(x), \dots, U_n(x)) = W(\varphi_1 \circ U_1(x), \dots, \varphi_n \circ U_n(x))$  for all possible  $\varphi_1, \dots, \varphi_n$ . Here one and the same  $W$  is employed for the initial profile and all of its transforms. It can be checked that on this construal,  $W$  satisfies Single Profile Neutrality. Hence, from the single profile theorem, it is dictatorial.
- (2) Weaker invariance properties than (1), for example:

$$W(U_1(x), \dots, U_n(x)) = W(\varphi_1 \circ U_1(x), \dots, \varphi_n \circ U_n(x)) \text{ if } \varphi_1 = \dots = \varphi_n,$$

or:

$$W(U_1(x), \dots, U_n(x)) = W(\varphi_1 \circ U_1(x), \dots, \varphi_n \circ U_n(x)), \text{ if there are } a > 0 \text{ and } b_1, \dots, b_n \text{ such that } \varphi_i \circ U_i = aU_i + b_i \text{ for all } i.$$

Conceivably, there may be no invariance at all imposed on  $W$ . For this continuum of cases, Single Profile Neutrality never holds, and it is easy to exhibit examples fulfilling the other conditions as well as Single Profile Non-Dictatorship. Standard examples are the Rawlsian maximin, which satisfies the first restriction,<sup>29</sup> and utilitarianism, which satisfies the second.

- (3)  $W(U_1(x), \dots, U_n(x)) = W'(\varphi_1 \circ U_1(x), \dots, \varphi_n \circ U_n(x))$  for all sets of  $\varphi_1, \dots, \varphi_n$ , with  $W'$  being *defined* by this equation. In other words, there are not just one, but infinitely many  $W$  functions, one for each set of transforms, all of them delivering the same social preference and even the same numerical values. This is an invariance statement again, but widely different from those in (1) and (2). On this construal, the  $W$  functions do not satisfy Single-Profile Neutrality, and it is possible to find non-dictatorial examples to meet the remaining conditions.

Among the three conceptions, only (2) involves interpersonal comparisons of utility. The two examples in (2) correspond to familiar comparisons, i.e., those of utility levels (for the maximin) and of utility differences (for utilitarianism). By contrast, (1) and (3) deny interpersonal utility comparisons, but emphatically, *in distinctive ways*. Construal (3)

<sup>29</sup> However, the maximin just satisfies Weak Pareto, not the full strength of the Pareto Principle.

exactly formalizes Bergson's 1938 claim that the social welfare function depends on indifference loci alone, which amounts to denying any interpersonal utility comparisons. Construal (1) involves more than this denial. It also imposes *welfarism* in the following heavy form. Not only are utility data sufficient to determine the social preference, irrespective of the physical descriptions of the states, but the mapping from these utility data to the social preference is *fixed*, whether utility data are computed with  $(U_1, \dots, U_n)$  or any authorized transform. Very roughly speaking, Single Profile Neutrality may be decomposed into a denial of interpersonal utility comparisons and this further component, which I will refer to as *strong welfarism*. A crucial point, which Fleurbaey and Mongin (2005) spell out formally, is that the full force of Single Profile Neutrality, not only its denial of utility comparisons, is needed in order to derive dictatorship. The latter does not follow from (3) alone.

The previous taxonomy explains why Samuelson and his opponents wrote at cross-purposes throughout the controversy. The former interpreted the Bergson–Samuelson function in the light of (3) exclusively, while the latter theorists considered only (1) and (2). Unfortunately, neither side was sophisticated enough to realize that the other was conceiving of the  $W$  function in a way different from its own. The taxonomy serves also to clarify the various interpretations of  $W$  that Bergson and Samuelson broached simultaneously in their early work, and it helps locate Samuelson's intellectual change. By contrast with the 1947 *Foundations*, his 1977 paper pursues only one interpretation of  $W$ , which is (3).<sup>30</sup> The paper improves on the treatise in another respect. At long last, Samuelson offered a counterexample of a Bergson–Samuelson function that was not dictatorial (1977: 84–86; see also 1981: 234). This function fits all the social choice theorists' conditions except for Single Profile Neutrality, which is thus shown to be an extraneous addition to the new welfare economics. The social choice theorists ignored Samuelson's relevant reply probably because they were still concentrating on the other suggestions contained in his previous work. Accordingly, even after Samuelson's attempted clarification, they felt that the following dichotomy was compelling: *either the Bergson–Samuelson function is dictatorial (= (1)), or it makes interpersonal comparisons of utility (= (2))*. This has remained the received view of the controversy and – by a further retrospective simplification – of the founding debate between Arrow and Bergson.<sup>31</sup>

With *Collective Choice and Social Welfare* in 1970 and a series of related papers, Sen set the stage for a new style of normative economics in

<sup>30</sup> Statement (3) formalizes the passage of the *Foundations* (1947: 228) quoted in section 4.

<sup>31</sup> Parks (1976: 450) has a very clear statement of the dilemma, and it is reiterated in Kemp and Ng (1977), Roberts (1980: 449), Sen (1986: 1149, with qualifications), Hammond (1991: 226). However, the last writer takes a different stand in Fleurbaey and Hammond (2004).

which utility comparisons were the focus of attention. Many eminent theorists followed in his footsteps. Their common framework of analysis was welfarist in the strong sense.<sup>32</sup> In such a framework, the only way to avoid dictatorship is to allow for interpersonal comparisons of utility functions. Accordingly, the older economists' question of *whether* such comparisons should be made gave way to the more specialized one of finding *which* were the appropriate ones, given varying normative commitments towards distribution. This led to sophisticated comparisons between utility comparisons, and especially, to a famous parallel between the Rawlsian-like leximin and utilitarian rules. Important as this work was and still is, it proceeded from a premature rejection, and effectively a misrepresentation, of the new welfare economists' contribution.

### 5.3 Individualism and the tradition of the field

The last significant point made by the welfare economists, notably Little (1952) and Bergson (1954), is that the functions  $E$  or  $W$  should not be interpreted as expressing the society's ordering but only as an ordering relative to the society. But then, whose ordering is it? Arrow's opponents insisted that it must be *a person's*. The welfare economist, they claimed, is very much like a consultant. He counsels officials who are to make large-scale decisions. He also counsels ordinary citizens who are willing to employ him in order to decide, say, whether or not they will support a tax reform. Whichever is the case, the argument continues, welfare analysis relates to somebody like you and me, not to a nebulous collective entity. The individual client communicates his piecemeal evaluative judgments to the welfare economist, who will summarize them into an ordering. This conclusion is unproblematic because the usual rationality considerations apply here to concrete individuals and are normatively compelling at this level.

This forceful answer would seem to cut the ground under Arrow's feet, and actually preclude the development of social choice theory altogether. I am not aware of an explicit rebuttal in the literature, which makes it worthwhile to offer one here. One version of the argument is easy to reject because it involves a serious confusion about methodological individualism. The welfare economists claimed in effect that collective entities ("the community as such," Bergson 1954: 243) did not *exist*. But it has been argued, I think, convincingly that methodological individualism is not the thesis that collectives do not exist. It is rather the (weaker) thesis that they cannot be automatically endowed with well-defined aims or objectives. Methodological individualism is a way of allocating the

<sup>32</sup> See the work surveyed by d'Aspremont (1985), Sen (1986), Mongin and d'Aspremont (1998), Bossert and Weymark (2004).

burden of proof. When it comes to, say, firms or nations, the burden of proof is on whoever claims that there is such a thing as the firm's objective function, or the nation's long-term interests. From this cursory discussion, I conclude that methodological individualism *supports*, if anything at all, the programme of investigating the conditions under which collective objectives can be constructed from individual objectives as the relevant data. This, broadly speaking, is the programme of social choice theory.

Here is a further counter-argument. Even granting the welfare economists' premiss that the social welfare ordering is a person's ordering, there are difficulties for their position. It amounts to discarding all of Arrow's conditions but one, i.e., Social Ordering. *A priori*, the individual client may be of any ethical type. He might not even accept the Pareto Principle, which would stop the analysis at the level of *E* without *W* being derivable. But if this is the case, what rôle is left for welfare economists? They are reduced to the menial task of teaching their clients how to maximize an objective function under predetermined constraints, whatever this function and these constraints may be. Surely, welfare economists have a higher opinion of their field. They write as if they have *some* theory of what counts as a suitable social objective; in particular, they never seriously envisage *E* being other than Paretian. What leads them astray here may be the implicit assumption that to form an ordering from a client's data is a trivial step. To be true to the "economist as consultant" picture, they would have to take into account the *construction* of the social welfare objective. It is at this prior stage that their traditional commitments, such as the Pareto Principle, enter into the picture. But if the individual client scenario is so enriched, social choice theory becomes relevant. Arrow's conditions, or rather the corresponding single-profile conditions, become interesting *prima facie*. They may be dismissed at the end of the day, but there is now some sense in saying that they belong to theoretical welfare economics.<sup>33</sup>

The welfare economists' arguments relied not only on the two lines of argument which I have disposed of, but also on invoking *the tradition of their field*. For instance, in the same quoted passage, Bergson wrote: "I have thought here to make explicit that this follows simply from the very nature of the discipline" (1954: 247). For all I know, this remarkable declaration clashes with the history of the subject. Admittedly, the notion of the economist as counselling individuals was commonplace in pre-war economics. But I do not think that anybody at that time believed that the whole of welfare economics could be reorganized around this single theme, especially when counselling was construed as narrowly as it was

<sup>33</sup> Compare the argument of this paragraph with Arrow's discussion of *individual* distributional ethics ("the ethics of Primus") in his *Collected Economic Papers* (I: ch.3, 55–56).

in Bergson and Little.<sup>34</sup> There is ample evidence that: (a) in a number of cases, welfare economists did not have any counselling scenario in mind; (b) when they did, they were prepared to extend their notion of a client to the collective entity, whatever that meant for them;<sup>35</sup> (c) they were not taking social welfare orderings as given, but constructing them, at least coarsely or in outline.<sup>36</sup>

To summarize the point bluntly, the new welfare economics, in the Bergsonian-Samuelsonian formulation of a social welfare function, was groping after something like the social choice aggregation problem. Arrow puts it in this way: "Social choice theory was a child, if unwanted, of the Bergson-Samuelson social welfare function viewpoint" (1983: 26). By denying the fact, welfare economists reformulated their enterprise in a bizarre way, which could not enhance its prestige among general economists. This denial provided them with a convenient excuse for not offering a complete analysis of Arrow's impossibility theorem and ensuing work. Bergson's and Little's thinking about the theorem never went beyond the disorganized stage of their initial reactions in the 1950s. Not being hindered to the same extent by preconceptions about welfare economics, Samuelson ended up offering the best analyses of the theorem and its single-profile variant. Still, even his most sophisticated comments leave much to be desired, as subsections 5.1 and 5.2 have shown.

## 6. A WORD ON THE COMPENSATION PRINCIPLE

The Compensation Principle of the new welfare economics provides a link with social choice theory that has attracted more attention than the Arrow-Bergson connection. However, it is conceptually less significant than the latter for a reason that needs spelling out. The critique of the Compensation Principle does *not* have to rely on using the impossibility theorem, unlike the critique of Bergsonian welfare economics, which absolutely requires it.

As is well known, the compensation tests attempted to extend the range of welfare judgments permitted by the Pareto Principle by taking into account the possibility of the gainers' compensating the losers. The Kaldor-Hicks test was inconsistent in that it led to cycles, actually obvious cycles of order two, but Scitovsky claimed that his more sophisticated "double test" would remedy this defect. Arrow argued that the Scitovsky test was also inconsistent. The logical skeleton of his refutation is this. The binary relation implied by the Scitovsky test is incomplete; a natural way

<sup>34</sup> Robbins (1932) might be an exception. But he is not a welfare economist, and his positions were often rejected by the new welfare economists for being too sketchy and too extreme.

<sup>35</sup> Evidence for (b) can be found in Lange (1942), and even more clearly in the debate over the second welfare theorem and the economic theory of socialism.

<sup>36</sup> Clear evidence for (c) can be found even in Bergson (1938: 323).

to make it complete is to declare two states  $x$  and  $y$  indifferent with each other if the test is conclusive neither for  $x$  against  $y$ , nor for  $y$  against  $x$ . However, indifference defined that way turns out to be intransitive, as a three-alternative example demonstrates (1963: 45). This fairly straightforward piece of reasoning stands by itself, regardless of the impossibility theorem.

If Arrow had tried to base his refutation on the theorem, he would have said in essence the following. Take *any* binary relation  $R$  that extends the partial ordering implied by the Pareto Principle and makes it complete. If  $R$  results from a social welfare function, then assuming the Arrovian conditions other than Social Ordering, one must conclude that  $R$  is intransitive. This sounds like a powerful critique because, in contrast to the previous argument, it does not depend on the particular way of making the Scitovsky relation complete. It does not even depend on selecting the Scitovsky relation in the first instance, and thus seems to provide an impressive refutation of the Compensation Principle *per se*. However, the argument requires one to apply Independence of Irrelevant Alternatives to the Arrow function that formalizes the Compensation Principle, and here we stumble on the same difficulties that have been spelled out for the Bergson–Samuelson social welfare function. It is difficult to decide whether this substantial objection crossed Arrow's mind when he decided to refute the Scitovsky test by means of an example rather than the impossibility theorem. Rhetoric expediency might have been the conscious reason: Scitovsky's simple idea called for an equally simple rebuttal.

It is instructive to compare the two arguments envisaged here with Chipman and Moore's (1978) detailed refutation. These authors establish that each test, including Scitovsky's, is cyclical by constructing general equilibrium positions. Arrow's numerical example and the suggested refutation through the impossibility theorem deliver the same conclusion without attempting to satisfy this economically relevant constraint on the set of social states. Chipman and Moore's argument is more telling, but it is also more remote from social choice theory. It is disappointing to conclude that the Compensation Principle does not fit in with the present discussion of progress in normative economics.

## 7. SOCIAL CHOICE THEORY AND THE CONDITIONS OF PROGRESS

I return now to the abstract criterion of progress by relating it to the main case study. In section 5, the word "problem" has come to mean two different things. I argued that the *general* problem of aggregating individual utility functions was part of the conceptual background of the new welfare economics, even if its spokesmen did not recognize it. Besides, there was the *specific* problem created by the impossibility theorem, which was of course invented by social choice theory, but must also count as a

problem for the new welfare economics, given that the general problem was in the air. I will discuss the specific problem exclusively because the general problem is too vague to permit precise comparisons with the four requirements.

Social choice theorists formulated the specific problem under the guise of a dichotomy – either the Bergson–Samuelson function is dictatorial or it involves interpersonal comparisons of utility – which made the choice of solution obvious. They said that interpersonal comparisons of utility were unavoidable in general, and then proceeded to prioritize some specific ways of making these comparisons. This would constitute evidence of both a problem and a solution meeting requirement (1) *if the dichotomy were compelling*. But section 5 has shown that this was not the case. Given the third possibility that social choice theorists omitted – to reject the strong welfarism component of Single Profile Neutrality, or, using Arrow's list, Independence of Irrelevant Alternatives –, their problem was inadequately formulated and their solution was at best optional.

Despite this persuasive objection I will argue that (1) is fulfilled after all. My argument depends on a crucial move – in order to describe the specific problem for T and its solution in T', I propose to adopt not the perspective of T' at the time of the controversy between T and T', but *today's* perspective on both theories, thus taking full benefit of hindsight. We now understand Arrow's impossibility theorem and its single profile variant much better than in the 1970s. The specific problem for the new welfare economics has been shown to have a trilemma structure, and its solution accordingly to involve two possibilities, i.e., to make interpersonal comparisons of either utility values or other individual data (the latter is equivalent to rejecting strong welfarism).<sup>37</sup> I will be able to conclude that requirement (1) is met if I manage to argue, first, that this up-to-date formulation of the problem belongs to social choice theory, not welfare economics or any other theory, and second, that the corresponding solution is forthcoming in social choice theory, not elsewhere. The first point is easy to defend. To restate the problem required one to isolate what, in the assumptions of the impossibility theorem, went beyond the denial of interpersonal utility comparisons, given the other background conditions, *and it is social choice theory that permitted this analysis*. Fleurbaey and Mongin (2005) provide the desired decomposition of Single Profile Neutrality by relating it to Sen's concept of the "social welfare functional,"

<sup>37</sup> Samuelson's 1977 example of a well-behaved social welfare function illustrates the latter kind of individual comparisons. He fixes a ray through the origin in the Euclidean space of commodity baskets, and measures the individual assessments of  $x$  and  $y$  in terms of the distances between the indifference curves of  $x$  and  $y$  to the origin. Kemp and Ng (1977) mistook Samuelson's procedure for a comparison between cardinal utilities, a confusion that Mayston (1982) exposed without convincing them.



thus employing the same formal tool as the social choice theorists they criticize. Those few economists who carried out the same analysis also relied on social-choice-theoretic tools.<sup>38</sup> To draw an easy contrast, consider Samuelson again. Having never fully mastered the Arrow–Sen concept of a functional linkage between individual characteristics and social preferences, he was unable to generalize his 1977 counter-example. Even his late papers in 1981 and 1987 remain at the level of imperfectly analyzed particular cases. Also, he persistently misapprehended Independence of Irrelevant Alternatives by failing to connect it with Strong Neutrality. This stopped him a long way from reaching a proper formulation of the specific problem.

The second point is not so easy to argue as the first. In retrospect, I have found only one early piece in social choice theory that belongs to the unexplored line, i.e., Pazner and Schmeidler's (1978) article on the "equivalent-egalitarian" criterion, which in effect pursues Samuelson's unfinished 1977 analysis.<sup>39</sup> This work borders on the fourth stage of normative economics because of its implicit rejection of welfarism, and it is indeed the fourth stage which brought out its potential clearly. Other attempts to elaborate on the unexplored line are recent and quite clearly not limited to social choice theory.<sup>40</sup> I conclude that this theory is responsible for correctly stating the problem, but not the whole of its solution. In order to dispose of this complication, I will modify my tentative criterion of progress in the last section.

Both to support the argument that requirement (1) is met and to reach the same conclusion for (2) and (3), I will pause and clarify the sense in which, generally speaking, social choice theory can be said to *resolve* problems. Many in the field are concerned mostly with exploring the compatibility or otherwise of given normative assumptions, without taking sides strongly for or against them. They might point out that an assumption is apparently acceptable or open to criticism, but would normally refrain from entering a sustained normative debate. The problems they are interested in take as their data some list of "axiomatic" conditions and their solutions take the form of either an impossibility theorem (e.g., Arrow's five conditions are incompatible) or a positive

<sup>38</sup> Pazner (1979), Mayston (1982), Blackorby, Donaldson, and Weymark (1990), Fleurbaey and Maniquet (in particular 1996 and 1999).

<sup>39</sup> An allocation  $(x_1, \dots, x_n)$  is *egalitarian-equivalent* if there is a benchmark vector  $\bar{x}$  such that each individual  $i$  is indifferent between  $\bar{x}$  and his component  $x_i$  in the allocation. When  $\bar{x}$  refers to the total resources, Pareto-optimal egalitarian-equivalent allocations enter Samuelson's 1977 example as special cases. This is shown in Fleurbaey and Mongin (2005: section 3).

<sup>40</sup> Fleurbaey and Maniquet (see their 1999) have recently proposed solutions in the style of the unexplored line. They are expressly intended to bridge the gap with public economics, and if only for this reason, it would be implausible to locate them in social choice theory exclusively.

characterization (e.g., utilitarianism is characterized by such-and-such list). With this *formal* interpretation of its task, social choice theory cannot regard the latter class of results as being more important than the former. It is true that impossibility theorems call for further investigations, while positive characterizations sound definitive, but this very argument could be invoked to say that the former are deeper than the latter. Now, beside this, currently predominant, formal notion of problem-solving, there is another one, which makes the normative discussion a very substantial part of the social choice theorist's activity. For a significant minority group – which I would argue includes Arrow and Sen themselves – solutions should be given at the *substantial* level of normative decisions made for or against a given set of conditions, while the formal statements play the rôle of preliminary groundwork. With this more commonsensical interpretation of its task, social choice theory will deemphasize impossibility theorems; positive characterizations are what matter more.<sup>41</sup> After contrasting the two groups with each other, I hasten to add that they overlap massively in their ordinary work. Some contributions are clearly purely formal, others are clearly substantial or at least offered as such, but a good deal of the problem-solving activity in the field falls in between.

This sketch needs comparing with what we know of the new welfare economists' attitude towards normative commitments. They were wary of certain "value judgments" and willing to indulge in others. They took the Pareto Principle to be both normatively commendable and indispensable, and they regarded judgments of interpersonal comparisons as being both normatively dubious and dispensable. These two *substantial* commitments defined the range of acceptable problems for which solutions could be sought. Within this substantially predetermined range, solutions were mostly offered at the *formal* level, as is apparent in Bergson, Samuelson, Lange and their followers. Comparisons between the new welfare economics and social choice theory will not be difficult to implement if one is careful to limit them to problem-solving activities of a given type, either formal or substantial.

This warning puts into proper perspective both requirements (1) and (2). If one reads Bergson and Samuelson's function as an entirely formal concept, the problem it raises can be compared with the entirely formal solution "make interpersonal comparisons of some kind." If (I think more appropriately) one interprets the function substantially, it will have to include the denial of interpersonal utility comparisons, and the more substantial solution becomes relevant: "make interpersonal comparisons of a non-utility sort, e.g., of individual indifference curves." With this

<sup>41</sup> Mongin (2004) argues that the position of the second group is not only conceptually richer than that of the first, but also to a large extent *unavoidable*, given the semantic constraints which go with the use of normative predicates.

warning in mind, and possibly taking into account the contribution of public economics, it is straightforward to check that requirement (2) is fulfilled.

Similarly, both formal and substantial resolutions are appropriate when considering (3). This requirement is easily satisfied by mentioning the problems in “mathematical politics” that social choice theorists have both raised and solved, from the early revival of the theory of committees in the 1950s to the current attempts to combine a description of the political process with a market equilibrium analysis. These problems were outside the initial range of the new welfare economics, and not only outside its range as it was tactically redefined once Arrow’s theorem became known. It is fair to recall at this juncture that modern social choice theory results also from Black’s *Theory of Committees and Elections* (1958) and earlier work on the same topic. Alternatively, one could stay even closer to Arrow’s theorem and mention the variant proved by Gibbard (1973), a justly famous result which opened up a whole new area of work – i.e., the non-manipulability of social choice decisions.<sup>42</sup>

Given today’s wide deployment of normative economics, it is not difficult to argue that the asymmetry condition (4) is fulfilled. On the weak reading of this condition, it is satisfied if the new welfare economics fails to solve an unresolved problem of either social choice theory or public economics, *or* fails to solve some problems that these theories do resolve, *or* fails to raise and eventually solve a problem of its own that these theories are silent about. On the strong reading, all three failures would be required. The weak reading seems to be preferable; if not, the criterion of progress would very rarely apply. However, in the present instance, the failure is multiple, and there is no need to decide between the weak and strong reading.

## 8. THE ASSUMPTIONS OF WELFARE ECONOMICS AND THE FOURTH STAGE OF NORMATIVE ECONOMICS

Although the main point has already been argued, i.e., that the third stage was a progressive one, I would like to take a broader view of my topic and briefly reexamine the basic assumptions of welfare economics. As will become apparent, the point is to relate them to *current* work, i.e., the fourth

<sup>42</sup> A methodological dispute is likely to take place in connection with this and related examples. Some writers in normative economics (e.g., Fleurbaey 1996) believe that non-manipulability, and other implementation concepts, belong to an area different from normative economics. As they construe it, the latter is concerned solely with norms and evaluations, not with the way in which these can be achieved in the economy. It seems that normative economics must be concerned with implementation issues, if only because they count among the considerations weighing for or against evaluative criteria.

stage of normative economics. This will lead me further to clarify the sense in which the third stage was progressive.

Welfare economics relies on conceptually loaded assumptions that have become better and better understood, and more and more heatedly criticized, with the passing of time. The following list attempts to summarize them. I state them in terms of the ideal concept of normative economics that welfare economics is supposed to encapsulate.

- (I) Normative economics is an exclusively teleological theory. That is to say, it will select a notion of the social good, and it will make all its evaluations and derived prescriptions dependent on this chosen notion.
- (II) The chosen notion of social good is social welfare. Social welfare is initially an undefined term in normative economics. It will be explicated in terms of the next conditions.
- (III) Social welfare in any circumstances is entirely determined by the data of individual welfare given these circumstances, and it increases when these data show an increase in individual welfare. Normative economics makes this claim precise in terms of the Pareto Principle, as interpreted in welfare terms.
- (IV) Normative economics is not concerned with social states in general. Only economic variables enter its description of the states.<sup>43</sup> In effect, the economic variables to be taken into account are the quantities of commodities consumed and of factors supplied by the individuals. The commodities may be either private or public goods.
- (V) Individual welfare can be measured by an index of preference satisfaction.
- (VI) The index of preference satisfaction summarizes the individual's choice behaviour ("revealed preference theory").
- (VII) The index can be endowed with the standard properties of an ordinal utility function. Monotonicity or at least non-satiation is typically imposed, and sometimes convexity as well. Other assumptions will have to be introduced to deal with risk and uncertainty, and this is done again by borrowing standard microeconomic constructions, such as the von Neumann–Morgenstern utility function.
- (VIII) The index is not comparable from one individual to another.

This is a rough picture, but it is sufficient for the conceptual discussion.<sup>44</sup> Welfare economists generally do not distinguish (V) from

<sup>43</sup> This can be formally explicated by assuming that non-economic variables are *separable* from economic variables within each individual welfare function. This is not a light assumption to make.

<sup>44</sup> It has sometimes been said that welfare economics needed only to make assumptions about *variations* in individual and social welfare; see Little (1950). I discard this line of

(VI) because they take “revealed preference theory” for granted. So, the statement corresponding to (V) and (VI) jointly goes like this in welfare economics:

A person’s welfare map is defined to be identical with his preference map – which indicates how he would choose between different situations, if he were given the opportunity for choice. To say that his welfare would be higher in A than in B is thus no more than to say that he would choose A rather than B, if he were allowed to make the choice. (de Graaff 1957: 5)<sup>45</sup>

All of these assumptions can be, and indeed have been, called into question, either jointly or separately. Take (V) and (VI) together. Of course, welfare economists know that maximizing behaviour in the revealed preference sense does not have the same meaning as maximizing behaviour in the welfare sense. What they intend to say is only that the former can serve as a *measure* of the other for the purpose of the theory. Presumably, this is the reason why de Graaff employs the word “defined” in the previous quote. Then, domain considerations should come to the forefront. The (purely extensional) coincidence of the two kinds of behaviour can only be justified by appealing to the restrictive notion of social states in welfare economics. This means that we should really consider (V) and (VI) jointly with (IV). But even in this charitable reading, the claim is more than dubious. Suppose that I have to choose between various baskets of apples and bananas, a matter relevant to the “economic” notion of a social state. From the fact that  $x$  is my chosen basket, and  $y$  is not, the welfare economist still cannot infer that my welfare would be lower in  $y$  than it is in  $x$ . This is a *non-sequitur*. They may be all sorts of reasons why I choose  $x$  instead of  $y$ , not all of them have to do with my welfare. Quite trivially, my tastes for apples and bananas might induce me to choose a basket with, say, too many bananas for my welfare. Some will perhaps be tempted to reply that non-welfare reasons show up as violations of the consistency of choices, but this would be a gratuitous assumption to make. A more standard reply is this. One cannot say that I am choosing too many bananas for my welfare if I really *choose* to have this basket. But this is tantamount to saying that, after all, welfare *is* the same thing as choice – a claim that was discarded at the outset as implausible. Notice that the familiar contention, “people are the best judges of their own interest,” is not sufficient to warrant the conclusion that choices provide a measure of welfare. The claim may be true without the people’s good judgment surfacing in their actual choices.

analysis partly for simplicity, partly because it does not seem very plausible to investigate variations in a quantity without saying to what the quantity refers.

<sup>45</sup> Compare with related statements in Boadway and Bruce (1984: 8), Little (1950), Mishan (1969: 23–25).

One way or another, the arguments just sketched have been made several times.<sup>46</sup> What I want to stress in connection with the present analysis of progress is that this seemingly commonsensical critique has entered normative economics only recently. It is not well taken by social choice theory, which generally has little to contribute to *the interpretation* of the preference concept. For most social choice theorists, preferences are just preferences, whatever that means; and if they are pressed to provide an interpretation, they might very well follow the welfare economist into the trap of “defining” welfare by choice.<sup>47</sup> It is only in the work currently pursued about non-standard indexes of welfare, especially in connection with Sen’s (1985) “functionings” and “capabilities,” that the critique above has become broadly understood.

A different (and more sophisticated) critique of welfare economics results from focusing on (IV) and (V), while putting (VI) aside. To relate an economic notion of welfare to *any* concept of preference raises possible objections. Sen’s (e.g., 1979, 1985) arguments usually proceed by considering *actual* preferences – “tastes” in Arrow’s terminology. But it is possible to introduce a notion of *improved* preferences that is located somewhere between “tastes” and “values,” i.e., preferences for the individual’s own good.<sup>48</sup>

These issues are often discussed in connection with the already mentioned concept of *welfarism*. In Sen and others, it refers to the view that individual *utility* data are both necessary and sufficient to form an index of social welfare. This is also the definition employed thus far in this paper. It has the defect of trading on an unspecified notion of “utility,” and in the present context of conceptual discussion, it seems preferable to fix a more substantial conception of welfarism, as claiming that individual *welfare* data are both necessary and sufficient to form an index of social welfare. Then, welfarism becomes identical with assumption (III) in the list. The argument against sufficiency can be made in terms of socially undesirable aspirations, as in Hare’s (1976) fanatic example or in Sen’s (1970) Paretian Liberal paradox. The case against necessity is not so straightforward to argue, and might necessitate considering the pitfalls of the Pareto Principle in the uncertainty context, which would involve assumption (VII) in the discussion.<sup>49</sup> I skip the discussion of the more basic commitments (I) and (II), which Sen and his followers have also come to question. Roughly

<sup>46</sup> Some were already made by the philosophers (not the economists!) participating in the conference on Human Values and Economic Policy (1967). Further occurrences are, among others, Broome (1978), Sen (1985), Mongin and d’Aspremont (1998).

<sup>47</sup> This happened several times over at recent meetings of the Society for Social Choice and Welfare.

<sup>48</sup> This sense of preference is suggested by the important work of Griffin (1986) and Harsanyi (1977). Mongin and d’Aspremont (1998: 388–401) follow the same direction.

<sup>49</sup> On this line of objection see, e.g., Mongin and d’Aspremont (1998).

speaking, it involves either changing (II) to enlarge the notion of social good beyond that of social welfare,<sup>50</sup> or replacing in (I) “to form a notion of the social objective” by “to evaluate social states,” so as to make room for deontological considerations.<sup>51</sup>

This bird-eye review was meant to support two methodological claims. First, as already emphasized, it was only long after the early stages of social choice theory that the argument against the new welfare economics was properly sorted out. I mentioned Arrow’s occasional anticipation of a far-reaching critique of the new welfare economics, i.e., a critique which would hit not only the Bergsonian Economic Welfare Function, but the Paretian core of welfare economics. Whatever Arrow’s intentions were in 1951 and in 1963, I do not think that he had the conceptual means of pursuing such a critique. The current discussions help to formulate it more appropriately. Second, there is a claim that is in a sense reciprocal to the previous one. The current discussions are usefully reorganized within the framework of a step-by-step refutation of the new welfare economics – even though the latter is old hat for many of today’s readers. Precisely because they embody an intermediary stage of critical thinking, the Arrovian and post-Arovian theories of the 1950–1980s are *not* a good polemical target to choose for “post-welfarist” writers. It is better to shoot at a theory which is blunter about its conceptual commitments.

This brief excursion into the fourth stage illustrates a relevant generality about the pace of progress in normative economics. Not only has this pace proved to be painfully slow, but it appears to follow a lag pattern. The most important semantic findings about the second stage are becoming available only now that normative economics has entered its fourth stage. In a rough parallel, section 7 has argued that the *logical* problem surrounding the Bergson–Samuelson function has been sorted out only recently, and that the fourth stage perspective inspired some of its solutions. *Progress in normative economics can be appreciated only by comparing non-successive theories.* Such lags are perhaps not surprising given the problem-based criterion adopted for assessing progress. Indeed, problems have a life of their own, some of them being quickly clarified, others dragging on for years.

## 9. CONCLUSIONS AND QUALIFICATIONS

By way of conclusion, I return to the tentative definition of progress, and discuss qualifications and refinements of the four requirements. To echo the last comment, I propose to reformulate (1) as follows: T’ points out

<sup>50</sup> Presumably, the work on “capabilities” follows this line of teleological, non-welfarist thinking.

<sup>51</sup> As in the work on rights stemming from another part of Sen’s work (e.g., 1981).

a problem that is unresolved in  $T$ , and this problem is resolved by  $T'$  either alone or in collaboration with some  $T''$  succeeding  $T'$ . This generalization seems to be unproblematic, given that the criterion aims only at making the commonsense notion of progress more precise, and it is intuitive that a theory is progressive with respect to another, not only if it contains the full clarification and resolution of a problem raised by the latter, but also if it prepares this final stage significantly. It would be interesting to collect scientific cases of the lagged manifestations of progress, but the growth-of-knowledge literature seems to have generally neglected this possibility.

I have already said that the unresolved problem in (1) may be like an anomaly for  $T$  or like a novel fact pointed out by  $T'$ . Boldly generalizing on the irrelevance of Compensation Tests to the present inquiry, I tend to believe that cases of the first kind will be scarcer than those of the second kind in normative disciplines. This means that the assessment in these disciplines will typically be controversial, since one must expect the  $T$  theorists to deny what the  $T'$  theorists claim, i.e., that there is a problem for  $T$ . The standard philosophy-of-science suggestion to decide between the two camps would be to resort to an *external decision procedure*, and the latter would go roughly like this. Investigate the formal languages of  $T$  and  $T'$ , as well as the intended interpretations of sentences produced by  $T$  and  $T'$  using their respective languages. If the problem made explicit by  $T'$  with its own syntactical and semantic resources could *also* have been formulated in the language of  $T$ , and if once so formulated, this problem would have fallen within the range of interpretations intended by  $T$ , then you may conclude that it was a theoretical problem for  $T$ ; otherwise, the problem belongs to  $T'$  exclusively. To some extent this abstract description fits the case study. At least, this is how I started discussing Arrow's conditions and whether they apply to the new welfare economics. However, once confronted with a crucial, but syntactically and semantically equivocal expression, "the Bergson–Samuelson social welfare function," I had to enter the economists' "conversations," as the fashionable slogan goes. Samuelson ended up restricting the initial sense of his welfare function, and one may wonder whether this shift was not in part the result of the social choice theorists' intervening work, in the same way as Bergson and Little altered the scope of welfare economics in reaction to Arrow's criticisms. This is a case where the semantics and even the syntax of  $T$  are adjusted after the fact to those of  $T'$ , and it prevents the procedure from delivering any relevant information, since the  $T$  theorists blur comparisons with  $T'$ . I would expect such disturbing phenomena often to take place in normative disciplines. The only way to disentangle them is to subject the external decision procedure to a *pragmatic and rhetoric analysis*, as I did sketchily here and more thoroughly in Fleurbaey and Mongin (2005).



This said, my approach is not pragmatic or rhetorical *from the start* contrary to McCloskey's (1994) approach or Dascal's (1998) more recent alternative.<sup>52</sup> Rather, it *complements* the syntactical and semantic analysis when they prove to be dubious or inconclusive. How this works can be illustrated by the above discussion of Arrow's unfinished attack against Bergson. Arrow had full control over the syntactical and semantic parts of his argument, but he chose to argue directly for Independence of Irrelevant Alternatives instead of trying to show that this condition was included in his opponent's position. A pragmatic (communication) disaster followed from this initial rhetoric move: for a long time, social choice theorists believed the condition to be unproblematic when they compared Arrow and Bergson. Here, pragmatic and rhetoric considerations come to the rescue of syntax and semantics, which would not provide the full picture.

I move on to requirement (2), which I propose to qualify in the same way as I did (1): *T' or subsequent theories T''* provide solutions to the main problems that T had already addressed and resolved in its own way. There is a persisting difficulty with the condition that T' or T'' solve *the main problems* of T, instead of *all* the problems of T. I said that the standard requirement is too strong to make the methodology of research programmes really applicable, but I must concede that my weakening is not only vague but even possibly inconsistent. It is conceivable that T' solve the main problems of T, while T'' solve the main problems of T', but not those of T. In this case, the "more progressive than" relation would be intransitive, which sounds absurd. Thus far, I have found no way out of this unpleasant dilemma.

Here is another less apparent difficulty for (2). The requirement that T' or T'' continue to solve the main problems of T is strong enough to ensure continuity, but not to exclude that dubious resolutions will be perpetuated. In the empirical sciences the corresponding requisite – roughly, that T' recovers most of the *corroborated* content of T – ensures in principle that what is common to T and T' is also what is valuable. Of course, the contrast must not be overdone: corroboration is arguably never definitive, and some problem resolutions can be. But there remains a substantial dissimilarity, and it might indicate that only progress "in the small," not progress "in the large" as in grand science, is really feasible for normative disciplines. Given the conceptual difficulties – or rather, the mass of confusions – which social choice theory unconsciously borrowed from the new welfare economics, the progress from one to the other is more limited than even my already qualified account suggests.

Concerning (3), I will only mention that this condition does not insist on originality, at least in the following sense. It is sufficient if traditional

<sup>52</sup> Dascal (1998) has recently proposed a method of studying scientific controversies and he has already illustrated it in economics in Dascal and Cremaschi (1999).

conceptions are *brought by T' to bear on the given problem*. The way in which public economics has repeatedly dragged the time-honored rule of utilitarianism into welfare discussions is a case in point. There is an analogy between the claim made here about originality and a view that surfaced in earlier philosophy-of-science discussions of novel facts. Against Lakatos's "temporal" view of evidence, it was argued – successfully, I believe – that a new theory could be corroborated by evidence already known before that theory came into existence.<sup>53</sup>

Requirement (4) was said to be easily met, a feature which makes the present study perhaps unrepresentative of economics generally. Outside normative economics proper, the field abounds in cases in which the first three conditions would apply more or less plausibly, but the fourth one would not be met. Consider the recent "non-expected" utility theories of risky choice. They solve a number of problems – some of them empirical, others normative – that were left open in von Neumann–Morgenstern theory. However, the smooth analysis of risk-attitudes provided by the latter has not found a full counterpart in the former. There are simple questions relative to insurance coverage or portfolio diversification that they cannot answer well. Perhaps they will do so in due course, but given their current state, one could argue that they are progressive only by adopting the weak (disjunctive) interpretation of (4).<sup>54</sup>

There is a warning I should finally make, lest the contribution of this paper be misunderstood. Welfare economics died, or rather disintegrated progressively, for many different reasons, *not all of which are connected with the emergence of a progressive alternative theory*. The pre-war controversy on market socialism could not be resolved by means of the fundamental welfare theorems, nor more generally in terms of the existing welfare economics; this was perhaps the first serious warning about its limitations. The post-war years witnessed an increasing discontent with its policy conclusions, not only because of the pervasiveness of externalities, but also because "second-best" considerations rose to the forefront. So the achievements of the new welfare economics proved dubious even to those who were *not* impressed by Arrow and his new style of theorizing. This suggests that one should be clear about the following distinction. There is a difference between claiming that the four conditions apply with some dose of success to the historical development of normative economics, and claiming that these conditions state the *causal factors* accounting for this development. The rational reconstruction of normative economics I have attempted is itself evaluative, and does not by itself purport to make

<sup>53</sup> See Zahar (1983) and Worrall (1985).

<sup>54</sup> The putative example of progress constituted by non-expected utility theories was discussed by Mongin (1988) along Lakatosian lines; it has been usefully re-examined by Guala (2000).

causality claims. But it suggests links that could possibly be turned into causality claims, and it is left for the historian to decide on that remaining issue.

## REFERENCES

- Archibald, G. C. 1959. Welfare economics, ethics, and essentialism. *Economica* 26:316–27
- Arrow, K. J. 1951. *Social choice and individual values*. Yale University Press; 2nd revised edition, 1963 (references are to the 2nd edition)
- Arrow, K. J. 1983. Contributions to welfare economics. In E. Cary Brown and R. M. Solow, *Paul Samuelson and modern economic theory*. McGraw Hill: 15–29
- Arrow, K. J. 1983–84. *Collected works*. Vol. 1: *Social choice and justice*. Vol. 2: *Social choice and justice*. Blackwell
- d'Aspremont, C. 1985. Axioms for social welfare orderings. In L. Hurwicz, D. Schmeidler, and H. Sonnenschein, eds., *Social goals and organization*. Cambridge University Press, 19–76
- Baumol, W. J. 1952. Review of *social choice and individual values*. *Econometrica* 20:110–11
- Bergson (Burk), A. 1938. A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics* 52:310–34
- Bergson, A. 1954. On the concept of social welfare. *Quarterly Journal of Economics* 68:233–52
- Bergson, A. 1966. *Essays in normative economics*. Harvard University Press
- Bergson, A. 1976. Social choice and welfare economics. *Journal of Public Economics* 171–90
- Black, D. 1958. *The theory of committees and elections*. Cambridge University Press
- Blackorby, C., D. Donaldson, and J. Weymark, 1990. A welfarist proof of Arrow's theorem. *Recherches économiques de Louvain* 56:259–86
- Boadway, R. W. and N. Bruce 1984. *Welfare economics*. Blackwell
- Bossert, W. and J. Weymark 2004. Utility in social choice. In S. Barberà, P. Hammond, and C. Seidl eds., *Handbook of utility theory*. Kluwer
- Broome, J. 1978. Choice and value in economics. *Oxford Economic Papers* 30:313–33
- Chipman, J. S. and J. Moore 1978. The new welfare economics 1939–1974. *International Economic Review* 19:547–84
- Dascal, M. 1998. The study of controversies and the theory and history of science. *Science in Context* 11:147–54
- Dascal, M. and S. Cremaschi 1999. The Malthus–Ricardo correspondence: sequential structure, argumentative patterns, and rationality. *Journal of Pragmatics* 31:1129–172
- Drèze, J. 1995. Forty years of public economics: a personal perspective. *Journal of Economic Perspectives* 9:111–30
- Feldman, A. M. 1980. *Welfare economics and social choice theory*. Kluwer
- Fleurbaey, M. 1996. *Théories économiques de la justice*. Economica
- Fleurbaey, M. and P. J. Hammond 2004. Interpersonally comparable utility. In S. Barberà, P. J. Hammond, and C. Seidl, eds., *Handbook of utility theory*, 2. Kluwer
- Fleurbaey, M. and F. Maniquet 1996. Utilitarianism versus fairness in welfare economics. Forthcoming in M. Salles and J. A. Weymark, eds., *Justice, political liberalism and utilitarianism: themes from Harsanyi and Rawls*. Cambridge University Press
- Fleurbaey, M. and F. Maniquet 1999. Compensation and responsibility. Forthcoming in K. J. Arrow, K. Suzumura, and A. K. Sen, eds., *Handbook of social choice theory*. North Holland
- Fleurbaey, M. and P. Mongin 2005. The news of the death of welfare economics is very exaggerated. To appear in *Social Choice and Welfare*
- Gibbard, A. F. 1973. Manipulation of voting schemes: a general result. *Econometrica* 41:587–601
- Graaff, J. de Van 1957. *Theoretical welfare economics*. Cambridge University Press
- Griffin, J. 1986. *Well-Being*. Clarendon Press

- Guala, F. 2000. The logic of normative falsification: rationality and experiments in decision theory. *Journal of Economic Methodology* 7:59–93
- Hammond, P. J. 1990. Theoretical progress in public economics: a provocative assessment. *Oxford Economic Papers* 42:6–33
- Hammond, P. J. 1991. Interpersonal comparisons of utility: why and how they are and should be made. In J. Elster and J. E. Roemer, eds., *Interpersonal comparisons of utility*. Cambridge University Press: 200–54
- Hare, R. M. 1976. Ethical theory and utilitarianism. In H. D. Lewis, ed., *Contemporary British philosophy*: 23–38
- Harsanyi, J. C. 1977. *Rational behavior and bargaining equilibrium in games and social situations*. Cambridge University Press
- Hook, S. 1967 ed. *Human values and economic policy*. New York University Press
- Howson, C. 1979. Methodology in non-empirical disciplines. In G. Radnitzky and G. Andersson, eds., *The structure and development of science*. Reidel: 257–66
- Kalinowski, G. 1969. *La querelle de la science normative*. Librairie Générale de Droit et de Jurisprudence
- Kemp, M. C. and Y. N. Ng 1976. On the existence of social welfare functions, social orderings and social decision functions. *Economica* 43:59–66
- Kemp, M. C. and Y. K. Ng 1977. More on social welfare functions: the incompatibility of individualism and ordinalism. *Economica* 44:89–90
- Keynes, J. N. 1890. *The scope and method of political economy* (4th edn, 1917). Kelley Reprints of Economic Classics (1963)
- Kitcher, P. 1993. *The advancement of science*. Oxford University Press
- Lakatos, I. 1963–64. Proofs and refutations. *The British Journal for the Philosophy of Science* 14:1–25, 120–39, 221–45, 296–342
- Lakatos, I. 1970. Falsification and the methodology of scientific research programmes. In I. Lakatos and A. E. Musgrave eds., *Criticism and the growth of knowledge*. Cambridge University Press: 91–196
- Lange, O. 1942. The foundations of welfare economics. *Econometrica* 10:215–28
- Laudan, L. 1977. *Progress and its problems*. University of California Press
- Le Breton, M. and J. A. Weymark 1996. An introduction to Arrowian social welfare functions on economic and political domains. In N. J. Schofield, *Collective decision-making: social choice and political economy*. Kluwer: 25–61
- Little, I. M. 1950. *A critique of welfare economics*. Clarendon
- Little, I. M. 1952. Social choice and individual values. *Journal of Political Economy* 60:422–32
- McCloskey, D. 1994. *Knowledge and persuasion in economics*. Cambridge University Press
- Mayston, D. J. 1982. The generation of a social welfare function under ordinal preferences. *Mathematical Social Sciences* 3:109–29
- Mishan, E. J. 1969. *Welfare economics: an assessment*. North Holland
- Mongin, P. 1988. Problèmes de Duhem en théorie de l'utilité espérée. *Fundamenta Scientiae* 9:289–317
- Mongin, P. 2004. Value judgments and value neutrality in economics. Cahier du Laboratoire d'économétrie, C.N.R.S & Ecole Polytechnique
- Mongin P. 2002. Is there progress in normative economics? In S. Boehm, C. Gehrke, and H.D. Kurz, eds., *Is there progress in economics? Knowledge, truth and the history of economic thought*. Edward Elgar
- Mongin, P. and C. d'Aspremont 1998. Utility theory and ethics. In S. Barberà, P. Hammond, and C. Seidl, eds., *Handbook of utility theory*, 1. Kluwer: 371–481
- Myint, H. 1948. *Theories of welfare economics*. Published for the London School of Economics by A. M. Kelley
- Parks, R. P. 1976. An impossibility theorem for fixed preferences: a dictatorial Bergson-Samuelson welfare function. *Review of Economic Studies* 43:447–50

- Pazner E. 1979. Equity, nonfeasible alternatives and social choice: a reconsideration of the concept of social welfare. In J. J. Laffont, ed., *Aggregation and revelation of preferences*. North-Holland
- Pazner E. and D. Schmeidler 1978. Egalitarian equivalent allocations: a new concept of economic equity. *Quarterly Journal of Economics* 92:671–87
- Pigou, A. C. 1920. *The economics of welfare*. Macmillan, 4th revised edn, 1932
- Pollak, R. A. 1979. Bergson-Samuelson social welfare functions and the theory of social choice. *Quarterly Journal of Economics* 93:73–90
- Popper, K. R. 1963. *Conjectures and refutations*. Routledge
- Robbins, L. 1932. *An essay on the nature and significance of economics*. Macmillan; 2nd revised edn, 1937
- Roberts, K. W. S. 1980. Social choice theory: the single and multi-profile approaches. *Review of Economic Studies* 47:441–50
- Samuelson, P. A. 1947. *The foundations of economic analysis*. Harvard University Press
- Samuelson, P. A. 1967. Arrow's mathematical politics. In S. Hook, ed., *Human values and economic policy*. New York University Press
- Samuelson, P. A. 1977. Reaffirming the existence of "reasonable" Bergson-Samuelson social welfare functions. *Economica* 44:81–88
- Samuelson, P. A. 1981. Bergsonian welfare economics. in S. Rosefield, ed., *Economic welfare and the economics of Soviet socialism: essays in honor of Abram Bergson*. Cambridge University Press: 223–66
- Samuelson P. A. 1987. Sparks from Arrow's anvil. In G. E. Feiwel, ed., *Arrow and the foundations of the theory of economic policy*. New York University Press
- Sen, A. K. 1970. *Collective choice and social welfare*. North Holland
- Sen, A. K. 1977. On weights and measures: informational constraints in social welfare analysis. *Econometrica* 45:1539–72. Reprinted in Sen (1982.), ch.11
- Sen, A. K. 1979. Utilitarianism and welfarism. *Journal of Philosophy* 76:463–89
- Sen, A. K. 1981. Rights and agency. *Philosophy and Public Affairs* 11:3–39
- Sen, A. K. 1982. *Choice, welfare and measurement*. Blackwell
- Sen, A. K. 1985. *Commodities and capabilities*. North Holland
- Sen, A. K. 1986. Social choice theory. In K. J. Arrow, M. D. Intrilligator, eds, *Handbook of mathematical economics*, III. North Holland: 1073–181
- Worrall, J. 1985. Scientific discovery and theory-confirmation. In J. C. Pitt, ed., *Change and progress in modern science*. Reidel: 301–31
- Zahar, E. 1983. Logic of discovery or psychology of invention? *British Journal for the Philosophy of Science* 34:243–61