The preference aggregation problem

Antonio Quesada[†]

Departament d'Economia, Universitat Rovira i Virgili, Avinguda de la Universitat 1, 43204 Reus, Spain

26th October 2008

Abstract
This paper presents the basic preference aggregation model and proves two of the most significant results in the theory of preference aggregation: Arrow's theorem and Sen's theorem.

Keywords: Social welfare function, preference aggregation, Pareto efficiency, independence of irrelevant alternatives, decisiveness, dictatorship.

JEL Classification: D71

[†] E-mail address: aqa@urv.cat. Financial support from the Spanish *Ministerio de Educación y Ciencia* under research project SEJ2007-67580-C02-01 and from the *Departament d'Universitats, Recerca i Societat de la Informació* (*Generalitat de Catalunya*) under research project 2005SGR-00949 is gratefully acknowledged.

1. Introduction

The distinctive feature of a microeconomic model consists of taking decisions made by individuals as the basic element of the model: in a microeconomic model, individuals and the way they make decisions are the building blocks of the model.

Individuals in a microeconomic model are typically assumed to make decisions on the basis of their preferences over the results that those decisions generate. Specifically, individuals are supposed to make decisions that lead to some of their most preferred results. So in game theory, the individuals (players) choose strategies in order to obtain the maximum payoff; in consumer theory, the individuals (consumers) choose bundles to maximize their utility functions; and in the theory of industrial organization, the individuals (firms) choose the amount of production to maximize their profit functions.

The view that every economic phenomenon involves individuals making some decision implies that all economic models are, in the last instance, microeconomic models. But even then the following question arises: is it legitimate to take the short cut consisting of dispensing with individuals in a model? That is, to what extent is satisfactory to have an economic model in which individuals are not explicitly modelled?

These considerations make the following question interesting: when can a collective be treated as an individual? For in cases in which a collective can be safely modelled as an individual then there is no need to deal explicitly with individuals: handling the collective directly, without having to care about individuals, would be a reasonably valid approximation for the extended model in which individuals are modelled. For instance, the models based on the idea of “representative consumer” analyze a group of individuals by just analyzing a single individual.

So there seems to be a legitimate research question to ascertain in which cases collectives can be treated as individuals. And as long as the basic information required from individuals are preferences, it appears necessary for collectives to be treated as individuals that collectives can be assigned preferences of the sort that can be assigned to individuals. The problem of when a collective can be considered an individual is in itself a decision problem, so economists have handled the question by means of microeconomic models. In particular, the standard model to study what type of preference can be ascribed to a collective takes the preferences of the individuals in the collective as the primitive element of the model. This paper presents two fundamental results obtained in this model.

2. Model

Let $N = \{1, \ldots , n\}$ be a non-empty finite set whose $n \geq 2$ members designate individuals. Let A be a finite set whose $m \geq 2$ elements represent outcomes, alternatives, choices or anything over which individuals can define preferences.

Definition 2.1. A preference on A consists of a binary relation p on A satisfying IRR, COM and TRA. The set of preferences that can be defined on A is denoted by L.

IRR. Irreflexivity. For all $x \in A$, it is not the case that $x\ p\ x$.
COM. Completeness. For all $x \in A$ and $y \in A \backslash \{x\}$, either $x\ p\ y$ or $y\ \ p\ x$.
TRA. Transitivity. For all $x \in A$, and $y \in A$ and $z \in A$, if $x\ p\ y$ and $y\ \ p\ z$ then $x\ p\ z$.

The interpretation of "$x\ p\ y$" is that the individual with preference p prefers x to y. A preference on A can be identified with a ranking (or linear order) of the elements of the set A. The definition of preference is restrictive because it does not allow indifference. This restriction is adopted for convenience, because the results in this paper can be extended to the case in which indifference is allowed.

Definition 2.2. A preference profile consists of an assignment of a preference on A to every individual in N. The set of preferences is denoted by L^n.

A preference profile P can be represented by an n-dimensional vector $(P_1, P_2, \ldots , P_n)$, where P_i stands for the preference that corresponds to individual $i \in N$ in preference profile P. Hence, a preference profile is a list of the preference held by every individual. In this model, the problem consists of associating, with every preference profile P, a preference P^* representing the collective preference when individuals have preferences as represented by P. Therefore, P^* can be viewed as a summarizing preference: when members of a collective have preferences represented by $P = (P_1, P_2, \ldots , P_n)$ then the collective can be attributed preference P^*. The rule creating the synthesizing (or collective) preference from the individual preferences is called social welfare function.

Definition 2.3. A social welfare function (SWF) is a mapping $f: L^n \rightarrow L$.

A SWF takes the individuals' preferences as inputs and outputs a collective preference associated with the group of individuals. A SWF is a mechanism transforming individual into collective preferences. Alternatively, a SWF attributes preferences to collectives taking into account the preferences of the members of the collective.

3. Axioms

There is an important presumption in the definition of SWF, namely, that the collective preference is drawn from the same set L as any individual preference. Such a fact makes a collective preference formally indistinguishable from an individual preference. After all, looking for a collective preference in the set L seems to be a reasonable requirement if one pretends a collective to be treated as an individual: if individual preferences satisfy IRR, COM and TRA, then it is a priori desirable to expect a collective preference to enjoy those properties.

This consideration suggests the idea that, when constructing the collective preference out of the individuals' preferences, a SWF should be "respectful" with the individuals' preferences. In fact, if a SWF disregarded the information contained in the individuals' preferences, why ask for those preferences?

There are many conditions that can be imposed on a SWF f to capture the idea that the SWF should respect the individuals' preferences. One is the Pareto principle. For non-empty subset $I \subseteq N$ of the set of individuals, $P \in L^n$, $x \in A$ and $y \in A \backslash \{x\}$, let $x \; P_I \; y$ abbreviate "for all $i \in I$, $x \; P_i \; y$".

PAR. Pareto principle. For all $P \in L^n$, $x \in A$ and $y \in A \backslash \{x\}$, $x \; P_I \; y$ implies $x \; f(P) \; y$.

PAR asserts that the process by means of which a SWF f generates the collective preference $f(P) \in L$ using the preference profile P as input should respect the unanimous preference of one alternative x over another alternative y. In other words, if all the individuals prefer x to y then, in the collective preference created by f, x should be preferred to y. This seems to be a plausible requirement: if all individuals prefer x to y, how could one contend that the collective does not prefer x to y?

Unfortunately, PAR is not always applicable: what happens if some individuals prefer x to y and the rest prefer y to x? One principle used in practice is majority: if a majority of individuals prefer x to y then the SWF should declare x collectively preferred to y. A political election can be viewed as a mechanism that implicitly constructs collective preferences: if political party x has more votes than party y and this more votes than z then it appears that one could infer that "society" prefers x to both y and z and also prefers y to z. It may be then tempting to demand a SWF to aggregate preferences using the majority rule.

Unfortunately, the Condorcet paradox (see Wikipedia (2008a)) shows this to be impossible: no SWF f exists such that, for all $P \in L^n$, $x \in A$ and $y \in A\backslash\{x\}$, if a majority of individuals prefers x to y the, in the collective preference $f(P)$, x is preferred to y. To see this, consider Example 3.1.

Example 3.1. Let $n = m = 3$, with sets $N = \{1, 2, 3\}$ and $A = \{x, y, z\}$. Define P to be the preference profile such that: (i) $x\ P_1\ y\ P_1\ z$; (ii) $y\ P_2\ z\ P_2\ x$; and (iii) $z\ P_3\ x\ P_3\ y$. Suppose that SWF f constructs the collective preference by applying the majority rule. Then, as a majority of individuals (1 and 3) prefer x to y, it must be that $x\ f(P)\ y$. And since a majority of individuals (1 and 2) prefer y to z, it must be that $y\ f(P)\ z$. Finally, given that a majority of individuals (2 and 3) prefer z to x, it must be that $z\ f(P)\ x$. Since $f(P)$ is a member of L, it satisfies COM. Therefore, $z\ f(P)\ x$ implies that not $x\ f(P)\ z$. And, as a result, TRA is violated: $x\ f(P)\ y$, $y\ f(P)\ z$ but not $x\ f(P)\ z$. This contradicts the fact that $f(P)$ belongs to L.

Example 3.1 shows that no SWF can be constructed using majority rule. The question is then how much of majority rule can be retained. Observe that PAR is one of the properties of majority rule. Another property of majority rule is that, when deciding whether x is collectively preferred to y, only the individuals' preference between x and y is taking into account. Hence, when determining the collective preference of x against y by majority rule the preference over the rest of alternatives is irrelevant. In Example 3.1, majority ranks x above y in the collective preference no matter how individuals rank z against x or against y. Consequently, majority rule satisfies the following property of independence of irrelevant alternatives.

IIA. Independence of irrelevant alternatives. For all $P \in L^n$, $Q \in L^n$, $x \in A$ and $y \in A\backslash\{x\}$, if, for all $i \in N$, $x\ P_i\ y \Leftrightarrow x\ Q_i\ y$ then $x\ f(P)\ y \Leftrightarrow x\ f(Q)\ y$.

By IIA, if each individual has the same preference over x and y in two preference profiles P and Q then, in both cases, the corresponding collective preferences $f(P)$ and $f(Q)$ have the same preference over x and y. IIA can also be seen as a form of respecting the individuals' preferences: if x was declared collectively preferred to y when some subset of individuals preferred x to y and the rest preferred y to x then, whenever the same subset of individuals prefer x to y and the rest prefer y to x, x has to be again declared collectively preferred to y.

For preference $p \in L$, let $p\big|_{\{x,y\}}$ represent the restriction of preference p to the set $\{x, y\}$. For instance, in Example 3.1, $P_1\big|_{\{x,y\}}$ is such that x is preferred to y whereas

$P_2|_{\{x,y\}}$ is such that y is preferred to x. With this notation, "for all $i \in N$, $x\ P_i\ y \Leftrightarrow x\ Q_i\ y$" means that, for each individual i, the restriction of i's preference to the set $\{x, y\}$ is the same in both preference profiles; that is, "for all $i \in N$, $x\ P_i\ y \Leftrightarrow x\ Q_i\ y$" is equivalent to "for all $i \in N$, $P_i|_{\{x,y\}} = Q_i|_{\{x,y\}}$": each individual has the same preference between x and y in both P and Q. Similarly, "$x\ f(P)\ y \Leftrightarrow x\ f(Q)\ y$" is equivalent to "$f(P)|_{\{x,y\}} = f(Q)|_{\{x,y\}}$". Consequently, after defining $P|_{\{x,y\}} = (P_1|_{\{x,y\}}, P_2|_{\{x,y\}}, \ldots, P_n|_{\{x,y\}})$, IIA can be alternatively expressed as follows.

IIA. For all $P \in L^n$, $Q \in L^n$, $x \in A$ and $y \in A\backslash\{x\}$, if $P|_{\{x,y\}} = Q|_{\{x,y\}}$ then $f(P)|_{\{x,y\}} = f(Q)|_{\{x,y\}}$.

Example 3.2. Let $n = 2$ and $m = 3$, with sets $N = \{1, 2\}$ and $A = \{x, y, z\}$. Let P and Q be the preference profiles such that $x\ P_1\ y\ P_1\ z$, $z\ P_2\ y\ P_2\ x$, $y\ Q_1\ x\ Q_1\ z$ and $x\ Q_2\ z\ Q_2\ y$. If SWF f satisfies IIA then f should rank y and z in the same way in both $f(P)$ and $f(Q)$, because: (i) 1 ranks y and z in the same way in both P_1 and Q_1 (1 prefers y to z); and (ii) 2 ranks y and z in the same way in both P_2 and Q_2 (2 prefers z to y). Note that IIA does not say that y must be preferred to z (or z to y) in both $f(P)$ and $f(Q)$. It just compels the SWF to make the same choice in $f(P)$ and $f(Q)$: the collective preference between y and z resulting in one case, must also result in the other case. On the other hand, for the preference between x and y, IIA imposes no restriction on f in this particular case, because $P_1|_{\{x,y\}} \neq Q_1|_{\{x,y\}}$.

4. Arrow's theorem

Definition 4.1. A non-empty subset I of individuals is decisive for $x \in A$ against $y \in A\backslash\{x\}$ in SWF f if, for all $P \in L^n$, $x\ P_I\ y$ implies $x\ f(P)\ y$.

That a subset I of individuals is decisive for x against y means that the group has the power to "impose" on the collective preference their unanimous preference for x against y. Thus, if I is decisive for x against y then, when all members of I prefer x to y, the collective preference dictates that x is preferred to y. It is worth noticing that PAR is equivalent to postulating that N is decisive for each $x \in A$ against each $y \in A\backslash\{x\}$.

Definition 4.2. A SWF f is dictatorial if there exists some individual $i \in N$ (called "dictator") such that, for all $x \in A$ and $y \in A\backslash\{x\}$, $\{i\}$ is decisive for x against y.

A SWF is dictatorial if there is some individual such that the collective preference always coincides with that individual's preference. That is, f is dictatorial if there $i \in I$ such that, for all $P \in L^n$, $x \in A$ and $y \in A \backslash \{x\}$, $x\, P_i\, y$ implies $x\, f(P)\, y$.

Theorem 4.3. (Arrow (1963, p. 97)). For $n \geq 2$ and $m \geq 3$, a SWF f satisfies PAR and IIA if, and only if, f is dictatorial.

Proof. "$\Leftarrow$" It is left as an exercise to verify that a dictatorial SWF satisfies PAR and IIA. "$\Rightarrow$" With $n \geq 2 < m$, let f be a SWF that satisfies PAR and IIA. The proof that f is dictatorial proceeds in two steps.

Step 1: for some $i \in N$, $x \in A$ and $y \in A \backslash \{x\}$, $\{i\}$ is decisive for x against y. This result will be proved by contradiction: if the negation of a sentence leads to a contradiction then the sentence must be true. So suppose otherwise: for no $i \in N$, $x \in A$ and $y \in A \backslash \{x\}$, $\{i\}$ is decisive for x against y. Equivalently, for all $i \in N$, $x \in A$ and $y \in A \backslash \{x\}$, $\{i\}$ is not decisive for x against y. As a consequence (why?),

$$\text{for all } i \in N, x \in A \text{ and } y \in A \backslash \{x\}, N \backslash \{i\} \text{ is decisive for } x \text{ against } y. \tag{1}$$

By the assumption that $m \geq 3$, choose any three different members x, y and z of A. Consider x and y. By (1), every group with $n - 1$ individuals is decisive for x against y. It will be shown that every group with $n - 2$ individuals is also decisive for x against y. To this end, let $P \in L^n$ be any preference profile in which the individuals' preferences restricted to $\{x, y, z\}$ are as follows.

P_1	P_2	P_3	P_4	…	P_n		$f(P)$
x	z	y	y		y		y
z	y	x	x		x	$\to$	x
y	x	z	z		z		z

It will be now argued that $f(P)$ restricted to $\{x, y, z\}$ is as indicated above. In P, all individuals except 1 prefer y to x. By (1), $N \backslash \{1\}$ is decisive for y against x and, accordingly, $y\, f(P)\, x$. Similarly, in P, all individuals except 2 prefer x to z. By (1), $N \backslash \{2\}$ is decisive for x against z and, as a result, $x\, f(P)\, z$. By TRA, $y\, f(P)\, x$ and $x\, f(P)\, z$ imply $y\, f(P)\, z$.

Consider the restriction $P|_{\{y,z\}}$ of the individuals' preferences to the set $\{y, z\}$, depicted next.

Individual	1	2	3	4	…	n		*Collective*
	z	z	y	y		y	$\rightarrow$	y
	y	y	z	z		z		z

By IIA, every profile $Q \in L^n$ with this preference configuration between y and z must result in a collective preference in which y is preferred to z. In other words, $N\backslash\{1, 2\}$ is decisive for y against z. Actually, IIA asserts that what happens once, happens always: since there is a situation (profile P) in which the above preference configuration between y and z leads to having y preferred to z, by IIA, whenever the above configuration occurs, y will be preferred to z. Let $Q \in L^n$ be any preference profile in which the individuals' preferences restricted to $\{x, y, z\}$ are as follows.

Q_1	Q_2	Q_3	Q_4	…	Q_n		$f(Q)$
z	z	y	y		y		y
x	x	z	z		z	$\rightarrow$	z
y	y	x	x		x		x

It will be now argued that $f(Q)$ restricted to $\{x, y, z\}$ is as depicted above. Observe that $P|_{\{y,z\}} = Q|_{\{y,z\}}$. As shown before, $y\ f(P)\ z$. Consequently, by IIA, $y\ f(Q)\ z$. On the other hand, $z\ Q_N\ x$: everybody prefers z to x. Hence, by PAR, $z\ f(Q)\ x$. Given $y\ f(Q)\ z$ and $z\ f(Q)\ x$, it follows from TRA that $y\ f(Q)\ x$. And IIA yields the result: all individuals except 1 and 2 (which have been arbitrarily chosen) prefer y to x and the SWF makes y collectively preferred to x. By IIA, this makes all groups of $n - 2$ individuals decisive for x against y.

The same line of reasoning can be applied to prove that all groups of $n - 3$ individuals are decisive for x against y, that all groups of $n - 4$ individuals are decisive for x against y, that all groups of $n - 5$ individuals decisive for x against y… and so on. It is evident that this line of reasoning will eventually contradict (1).

Step 2: if, for some $i \in N$, $x \in A$ and $y \in A\backslash\{x\}$, $\{i\}$ is decisive for x against y then, for all $x \in A$ and $y \in A\backslash\{x\}$, $\{i\}$ is decisive for x against y. To prove step 2, it suffices to show that $\{i\}$ decisive for x against y implies: (i) that $\{i\}$ is decisive for x against z; and (ii) that $\{i\}$ is decisive for z against y. So suppose $\{i\}$ is decisive for x against y. Choose $z \in A\backslash\{x, y\}$, which exists because $m \geq 3$. Case 1: $\{i\}$ is decisive for x against z. This means that, no matter the preferences of the rest of individuals over x and z, if i prefers x to z then the collective preference has x preferred to z.

Let $R \in L^n$ be any preference profile satisfying: (i) $x\ R_i\ y\ R_i\ z$; and (ii) for all $j \in N \setminus \{i\}$, $y\ R_i\ x$ and $y\ R_i\ z$. Notice that individuals different from i may have any arbitrary preference over x and z. Notice as well that, by TRA, $x\ R_i\ y$ and $y\ R_i\ z$ imply $x\ R_i\ z$.

The aim is to demonstrate that $x\ f(R)\ z$. First, by the assumption that $\{i\}$ is decisive for x against y, $x\ R_i\ y$ yields $x\ f(R)\ y$. Second, $y\ R_N\ z$ and PAR imply $y\ f(R)\ z$. And third, by TRA, it follows from $x\ f(R)\ y$ and $y\ f(R)\ z$ that $x\ f(R)\ z$. Given that $x\ R_i\ z$ has resulted in $x\ f(R)\ z$ no matter the preferences of the rest of individuals over x and z, IIA implies that $\{i\}$ is decisive for x against z.

Case 2: $\{i\}$ is decisive for z against y. Left as an exercise.■

5. Remarks

What is Arrow's theorem telling? The most evident lesson is that it is impossible to construct a SWF using two properties of the majority rule: unanimity (PAR) and requiring the collective preference over two alternatives to depend only on the individuals' preferences over these two alternatives (IIA). In this respect, Arrow's theorem generalizes the Condorcet paradox and makes evident the difficulties of trying to consider collectives as individuals by ascribing to collectives preferences that individuals could hold. For more on the interpretations of Arrow's theorem, see Wikipedia (2008b).

Arrow's theorem is a landmark in economic theory. For one thing, modern social choice theory emerged from this result; see Wikipedia (2008c) and Sen (1998). And, for another, Arrow's theorem is one of the single results in economic theory that has generated more literature; see, for instance, Kelly (2008). This literature has grown following two strands: one criticizing the result; and the other trying to escape from its negative conclusion (the existence of a dictator) by relaxing some of the assumptions.

Arrow's (1963, p. 97) theorem is, in fact, a more general result than the one presented here: it is still true when individuals, as well as the collective, can be indifferent between two alternatives. In that case, there also arises a dictator: if the dictator prefers x to y then the SWF declares x to be collectively preferred to y. The difference is that the collective preference is not identical with the dictator's preference: that the dictator is indifferent between two alternatives does not imply that the two alternatives are also indifferent in the collective preference.

Generalizations of Arrow's theorem have followed four basic routes; see Sen (1986) and Moulin (1994). One consisted of considering more general types of preferences as admissible; see, for instance, Blair and Pollak (1979) and references therein. Mas-Colell and Sonnenschein (1972) assume the arguably weakest type of preference that can be deemed admissible: acyclic preferences. A preference is acyclic if, for any sequence (x_1, x_2, … , x_{k-1}, x_k) of alternatives it is not the case that x_1 is preferred to x_2, x_2 is preferred to x_3, … , x_{k-1} is preferred to x_k and x_k is preferred to x_1. They postulate a condition MON of monotonocity: if x is preferred or indifferent to y in the collective preference then x becomes preferred to y if some individual just prefers x to y "a little more" (transforms indifference in preference for x or transforms preference for y in indifference).

Mas-Colell and Sonnenschein (1972, p. 189) Theorem 3 states that rules generating acyclic preferences and that satisfy PAR, IIA and MON have a weak dictator. A weak dictator is an individual i such that, for all $x \in A$ and $y \in A\backslash\{x\}$, if i prefers x to y then the collective preference cannot have y preferred to x (so x is preferred, or indifferent, to y). Results like this one have contributed to reinforce the view that Arrow's theorem is a robust result.

Another route tried to ascertain what would happen if PAR were removed. Wilson (1972, p. 484) proves that a SWF (as defined here) satisfying IIA is either dictatorial or inversely dictatorial, where SWF f is inversely dictatorial if there is $i \in N$ such that, for all $P \in L^n$, $x \in A$ and $y \in A\backslash\{x\}$, $x\ P_i\ y$ implies $y\ f(P)\ x$.

A third route considered domain restrictions. As defined, the domain of a SWF contains all the preference profiles. But it is well-known that Arrow's theorem fails on some restricted domains. For instance, majority rule creates collective preferences when preferences are single-peaked; see Black (1948) and Wikipedia (2008d). In addition, majority rule is consistent when there are only two alternatives, the case excluded by Arrow's theorem. Quesada (2002) identifies a subset of L^n in which PAR and IIA still suffice to generate dictators. For the simplest case in which Arrow's theorem holds (n = 2 and m = 3), 6 profiles are identified whose presence suffice for PAR and IIA to make the SWF dictatorial (compare them with the 13 profiles required in Feldman's (1980, ch. 10) proof).

But most effort has been devoted to ascertain the effects of weakening IIA. Since the literature is immense, just three contributions are singled out for the purpose of illustration. One is Baigent (1987), who slightly weakens IIA by requiring that if $P\big|_{\{x,y\}}$

$= Q|_{\{x,y\}}$ and x is preferred to y in $f(P)$ then x is preferred or indifferent to y in $f(Q)$. He shows that a SWF satisfying PAR and this weakening of IIA has a weak dictator.

A second interesting contribution is Denicolò (1998), who obtains a dictatorial SWF by replacing IIA with RID below. Define a non-empty subset I of individuals to enforce $x \in A$ against $y \in A\backslash\{x\}$ if for any preference profile $P \in L^n$ such that $x\ P_I\ y$ there is some profile $Q \in L^n$ satisfying $P|_{\{x,y\}} = Q|_{\{x,y\}}$ and $x\ f(Q)\ y$. This says that whenever all members of I prefer x to y there is some way of completing all the preferences so that x is collectively preferred to y.

RID. Relational independent decisiveness. For all $x \in A$, $y \in A\backslash\{x\}$ and $I \subseteq N$, if I can enforce x against y then I is decisive for x against y.

One of the most interesting recent contributions to the literature relaxing IIA is Campbell and Kelly (2000), who show how easy it is to weaken IIA and obtain non-dictatorial SWFs.

Turning to the criticisms to Arrow's theorem, it is worth mentioning Saari (1998), who severely criticizes IIA (and condition MIL in Sen's theorem); see also Saari (2001). In particular, he argues that IIA is an inappropriate requirement for SWFs because SWFs treat the transitivity of preferences as a valuable input and IIA is indifferent to the presence or absence of transitivity. His point is that preference aggregation procedures satisfying IIA do not take into account the transitivity of preferences and that, accordingly, it does not seem reasonable to impose on a procedure taking transitivity into account (SWFs) a property that neglects that information. He remarks (p. 255): "If we must build a vehicle using only oxen and carts, do not expect a Porsche. Similarly, if we can only use crude unsophisticated procedures, do not expect rational outcomes".

Quesada (2007) suggests an interpretation of SWFs that may lower the negative impact of Arrow's result. Arrow's theorem has been hailed as a negative and undesirable result because the outcome is the existence of a dictator, that is, an individual who has his preference coincide with the collective preference. But a dictator is something to worry about only in cases in which the dictator's preference is not "representative". The paper shows the sense in which a dictator can be seen to never have more "power" than 3 voters nor less than 2.

PAR and IIA can be viewed as "vertical" conditions, because they refer to whether some alternative is above another one in a preference ranking. Finally, Quesada (2003) and Houi (2006) have suggested a conceptual twist to the analysis of preference aggregation: to consider "horizontal" conditions instead, that is, conditions telling how to determine the alternative that occupies a certain position in the collective preference using as input the alternatives occupying that position in the individuals' preferences. Both authors show that positional conditions similar to PAR and IIA also generate dictatorial SWFs.

6. Sen's theorem

As discussed in Section 5, Arrow's theorem is father of many, many sons. One of the most reputed sons is Sen's theorem, which is motivated by the idea of how much power can a SWF ascribe to individuals. In particular, Sen (1970a, 1970b) is concerned with the possibility of having SWFs that are "liberal" in the sense that allows individuals to retain a minimum of freedom.

Specifically, Sen's motivating situation is the following: "Given other things in the society, if you prefer to have pink walls rather than white, then society should permit you to have this, even if a majority of the community would like to see your walls white". Therefore, if x = "individual i's walls are pink" and y = "individual i's walls are white" then, for i to enjoy a very weak form of individual liberty, i would have to be decisive for x against y and for y against x: if i prefers x to y (or y to x) then the collective preference should adopt i's view. Sen's theorem asserts that no SWF satisfying PAR can secure this weak form of individual liberty to two individuals.

Definition 6.1. A non-empty subset I of individuals is decisive over $\{x, y\}$ in SWF f if I is both decisive for x against y and decisive for y against x.

MIL. Minimal liberalism. There are $i \in I$, $j \in I\backslash\{i\}$, $x \in A$, $y \in A\backslash\{x\}$, $v \in A$ and $w \in A\backslash\{v\}$ such that $\{i\}$ is decisive over $\{x, y\}$ in f and $\{j\}$ is decisive over $\{v, w\}$ in f.

Theorem 6.2. (Sen (1970a)). For $n \geq 2$ and $m \geq 2$, there is no SWF f satisfying PAR and MIL.

Proof. Suppose f is a SWF satisfying PAR and MIL. Therefore, there are $i \in I$, $j \in I\backslash\{i\}$, $x \in A$, $y \in A\backslash\{x\}$, $v \in A$ and $w \in A\backslash\{v\}$ such that $\{i\}$ is decisive over $\{x, y\}$ in f and $\{j\}$ is decisive over $\{v, w\}$ in f. By definition of decisiveness, $\{x, y\} \neq \{v, w\}$: two individuals cannot be decisive over the same set. Despite this, there are two possibilities. Case 1: $\{x, y\} \cap \{v, w\} = \varnothing$. Consider any preference profile $P \in L^n$ in which the restriction of the individuals' preferences over the set $\{x, y, v, w\}$ are as follows.

P_i	P_j	rest of individuals			$f(P)$
w	y				x
x	v	w	y	$\rightarrow$	y
y	w	x	v		v
v	x				w

By the assumption that $\{i\}$ is decisive over $\{x, y\}$, $x\ P_i\ y$ implies $x\ f(P)\ y$. By PAR, $y\ P_N\ v$ implies $y\ f(P)\ v$. Consequently, by TRA, $x\ f(P)\ y$ and $y\ f(P)\ v$ imply $x\ f(P)\ v$. By the assumption that $\{j\}$ is decisive over $\{v, w\}$, $v\ P_j\ w$ implies $v\ f(P)\ w$. Thus, by TRA, $x\ f(P)\ v$ and $v\ f(P)\ w$ imply $x\ f(P)\ w$. By this contradicts PAR, because $w\ f(P)\ x$.

Case 2: $\{x, y\} \cap \{v, w\} \neq \varnothing$. Left as an exercise.■

Sen's theorem has also generated a large literature; see, by way of illustration, Wriglesworth (1985) and Wikipedia (2008e).

References

Arrow, K. (1963): *Social Choice and Individual Values*, 2nd ed., Wiley, New York.

Baigent, N. (1987): "Twitching weak dictators", Journal of Economics 47, 407–411.

Black, D. (1948): "On the rationale of group decision-making", Journal of Political Economy 56, 23–34.

Blair, D. and Pollak, R. (1979): "Collective rationality and dictatorship: The scope of the Arrow theorem", Journal of Economic Theory 21, 186–194.

Campbell, D. and Kelly, J. (2000): "Information and preference aggregation", Social Choice and Welfare 17, 3–24.

Denicolò, V. (1998): "Independent decisiveness and the Arrow theorem", Social Choice and Welfare 15, 563–566.

Houi, N. (2006): "Positional independence in preference aggregation: A remark", Social Choice and Welfare 27, 341–345.

Feldman, A. (1980): *Welfare Economics and Social Choice Theory*, Martinus Nijhoff, Boston.

Kelly, J. S. (2008): "Social choice bibliography", http://faculty.maxwell.syr.edu/jskelly/biblioho.htm, accessed the 27th October 2008.

Mas-Colell, A. and Sonnenschein, H. (1972): "General possibility theorems for group decisions", Review of Economic Studies 39, 185–192.

Moulin, H. (1994): "Social choice", in Aumann, R. and Hart, S. (eds.): *Handbook of Game Theory, volume 2*, North-Holland, Amsterdam, pp. 1091–1125.

Quesada, A. (2002): "Preference profiles sustaining Arrow's theorem", Economic Theory 20, 623–627.

Quesada, A. (2003): "Positional independence in preference aggregation", Social Choice and Welfare 20, 363–370.

Quesada, A. (2007): "1 dictator = 2 voters", Public Choice 130, 395–400.

Saari, D. (1998): "Connecting and resolving Sen's and Arrow's theorems", Social Choice and Welfare 15, 239–261.

Saari, D. (2001): *Decisions and Elections: Explaining the Unexpected*, Cambridge University Press, Cambridge.

Sen, A. (1970a): "The impossibility of a Paretian liberal", Journal of Political Economy 78, 152–157.

Sen, A. (1970b): *Collective Choice and Social Welfare*, Holden-Day, San Francisco.

Sen, A. (1986): "Social choice theory", in Arrow, K. and Intriligator, M. (eds.): *Handbook of Mathematical Economics, volume III*, North-Holland, Amsterdam, pp. 1073–1181.

Sen, A. (1998): Prize lecture for *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1998*, http://nobelprize.org/nobel_prizes/economics/laureates/1998/sen-lecture.pdf, accessed the 27th of October, 2008.

Wikipedia (2008a): http://en.wikipedia.org/wiki/Condorcet_paradox, accessed the 27th of October, 2008.

Wikipedia (2008b): http://en.wikipedia.org/wiki/Arrow%27s_impossibility_theorem, accessed the 27th of October, 2008.

Wikipedia (2008c): http://en.wikipedia.org/wiki/Social_choice, accessed the 27th of October, 2008.

Wikipedia (2008d): http://en.wikipedia.org/wiki/Majority_rule, accessed the 27th of October, 2008.

Wikipedia (2008e): http://en.wikipedia.org/wiki/Liberal_paradox, accessed the 27th of October, 2008.

Wilson, R. B. (1972): "Social choice without the Pareto principle", Journal of Economic Theory 5, 478– 486.

Wriglesworth, J. (1985): *Libertarian Conflicts in Social Choice*, Cambridge University Press, Cambridge.