Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 166

A Theorem on Preference Aggregation

Salvador Barberà

July, 2003

A THEOREM ON PREFERENCE AGGREGATION

Salvador Barberà*

Universitat Autònoma de Barcelona

Barcelona Economics WP nº 166

This Draft: July, 2003

Abstract

I present a general theorem on preference aggregation. This theorem implies, as corollaries, Arrow's Impossibility Theorem, Wilson's extension of Arrow's to non-Paretian aggregation rules, the Gibbard-Satterthwaite Theorem and Sen's result on the Impossibility of a Paretian Liberal.

The theorem shows that these classical results are not only similar, but actually share a common root.

The theorem expresses a simple but deep fact that transcends each of its particular applications: it expresses the tension between decentralizing the choice of aggregate into partial choices based on preferences over pairs of alternatives, and the need for some coordination in these decisions, so as to avoid contradictory recommendations.

* This paper is based on the second part of my invited lecture at the 5th meeting of the Social Choice and Welfare Society (Alicante, June 29 - July 2, 2000). I am grateful to Michael Florig, Jean Fraysse, Matthew Jackson, Cristina Rata, Peter Vida and especially to Carmen Beviá for their comments and help. I thank the editor of Social Choice and Welfare and one anonymous referee for their careful comments, and Bernard Monjardet for extensive historical remarks, from which I have drawn freely. Financial support from the Spanish Ministry of Education and Culture through grant PB98-0870, from the Spanish Ministry of Science and Technology through

1. Introduction

I present a general theorem on preference aggregation. The substance of the theorem is to exhibit the tension between two general principles. One principle would consider the possibility of decomposing the choice of a social aggregate into a set of partial decisions, each one based on pairwise comparisons of alternatives. The other principle would require that no agent plays a dominant role in the aggregation process. I show that, in the presence of some other requirements, decomposable rules can only avoid contradictory recommendations in the presence of the strongest coordination device: a single agent being determinant at all times on all aspects of the global decision.

The theorem implies, as corollaries, Arrow's Impossibility Theorem, Wilson's extension of Arrow's theorem to non-Paretian aggregation rules, the Gibbard-Satterthwaite Theorem and Sen's result on the Impossibility of a Paretian Liberal. Hence, one contribution of the paper is to provide a common framework that includes as particular cases all the different classes of aggregation rules that these classic theorems refer to. This allows us to understand the many connections that had already been noted among the substance and also among the proofs of these different results. In particular, it clarifies that the connections between Arrow's and Gibbard's results go deeper than any parallelism in their proofs; both come from a common root. It also allows us to note the intimate connection between other results, like Arrow's and Sen's, which had passed less noticed (with some significant exceptions that I refer to later).

Arrow's and Wilson's theorems consider social welfare functions. The Gibbard-Satterthwaite theorem applies to social choice functions, and Sen's result is expressed for social decision functions. All of these functions have a common domain: the profiles of individual preferences on some set of alternatives. All of them map these profiles into some type of related object, which we call an aggregate. Aggregates are different objects in each of these formulations. In Arrow's and Wilson's framework, an aggregate is a preference order satisfying the same properties that are required from individual preferences. In Sen's formulation, an aggregate is an acyclic binary relation (thus not necessarily transitive when social indifferences are allowed). In the context where the Gibbard-Satterthwaite theorem applies, aggregates are simply alternatives.

My theorem refers to preference aggregation rules. These are functions which map profiles of preferences on alternatives into some set of unspecified objects, to be called aggregates. They include all of the above as special cases. Specifically, the theorem concentrates on the class of preference aggregation rules that I call local. All of the classical results I have mentioned refer to special cases of local aggregation. It is thus important to clarify the concept immediately.

Local aggregation rules are based on the idea that different parts of the information contained in preference profiles can be used in making partial choices among aggregates, and that these partial choices eventually determine the one aggregate which corresponds to each

grant BEC2002-002130, and from the Generalitat of Catalonia through grant SGR2001-00162 is gratefully acknowledged.

profile. Not all preference aggregation rules admit a description in these terms, but many do. Formal definitions come in Section 2, but we should discuss immediately the meaning of local aggregation.

When given an aggregation rule, it may be hard to find out what is the algorithm, the procedure or the set of considerations which lead to associate a certain aggregate to a given preference profile. But here is one possible way to do it, which characterizes local aggregation: (1) First, identify some pieces of information that will be considered relevant. In our formulation of local rules, we'll consider that the set of comparisons made by the agents between each pair of alternatives are the pieces of relevant information. (2) Second, associate a set of aggregates to each piece of relevant information. We can consider this set as the collection of aggregates that are admissible, given the information at hand. For example, if aggregators are binary relations, (like those we obtain from majority rule), we can associate to each set of individual comparisons between x and y the set of all binary relations which rank x as preferred to y, or the set of all those which rank y over x. As another example, consider social choice functions, which aggregate preferences and choose one alternative to be the aggregate. In that case, a local rule could associate a set of alternatives to each piece of relevant information, and interpret that the elements in each of these sets are the potential aggregates whose choice is not precluded by the information available. (3) Third, describe the social aggregate as the result of combining the use of all the relevant pieces of information. Specifically, the chosen aggregate should be the one (and only one, in our formulation), which belongs to all the sets of non-excluded aggregates obtained from using all the available pieces of relevant information. They can be a binary relation, obtained as the intersection of sets of binary relations, or an alternative, resulting from intersecting sets of alternatives, or any other abstract object which belongs to the intersection of sets of objects in the relevant class of aggregates.

Limiting attention to local rules is somewhat restrictive. Many aggregation rules one can think of, like the Borda counts and other point voting rules, do not adjust to the structure required by this condition. But we shall see that a theorem on local rules is able to cover substantial ground.

More specifically, the theorem is about local and restricted aggregation rules. As already pointed out, the aggregates obtained from a local rule can be described as combinations of objects that satisfy some features, but not all combinations of features are necessarily acceptable to form an aggregate. For example, the feature of ranking x over y is acceptable for aggregates which are orders. So is ranking y over z. And so is ranking z over x. These three features would be compatible if the binary relation which results from aggregation was not restricted any further, as with majority rule. But the three features together cannot be met by an aggregate binary relation, if we further require this relation to satisfy transitivity, as Arrow or Wilson do, or acyclicity, as in Sen. Similarly, admissible features of an aggregate in the Gibbard-Satterthwaite setting are that the aggregate is not x, that it is not y, and so on, for any possible alternative. But the union of all these features cannot be met by any aggregate, because only the empty set, which is not an acceptable aggregate, can meet them all. Again in Section 2, I provide a formal definition for a local rule to be restricted, which is enough to capture all the limitations in the choice of an aggregate that is imposed by our classical reference theorems, in each of the cases.

As pointed out by Bernard Monjardet in private correspondence, this view of an aggregation process is not only natural but has a long tradition. I quote from Monjardet's generous comments:

"In his Memoire *Recherches sur la loi de croissance de I'homme* published in 1832 the statistician and social scientist Quetelet proposed to define the "mean man" of a population of men described by several measurable characteristics by taking the means on each characteristics. This conception was severely criticized by people like the mathematician Cournot which remarked (In *Exposition de la théorie des chances*, Paris 1843) that this mean man could be nothing else that an impossible man. Indeed one meets a similar "paradox" by considering the triangle obtained by taking the 3 means of the lengths of the corresponding sides of a set of rectangular triangles; this triangle is not necessary rectangular. More generally Cournot writes: "Lorsqu'on applique la détermination des moyennes aux diverses parties d'un système compliqué, il faut bien prendre garde que ces valeurs moyennes peuvent ne pas se convertir: en sorte que l'état du système dans lequel tous les elements prendraint à la fois le valeurs moyennes déterminées *séparément* pour chacun d'eux serait un état *impossible*"[1]. For Cournot, a 19 th century mathematician, a mean is any mathematical mean defined on numerical values. But more generally any p-ary operation defined on a set and satisfying some properties (like idempotence) can be called a mean and the Cournot text can be applied to any aggregation procedure which is "local" in the sense that it is obtained by

1) Decomposing the set of complex objects (to aggregate) into sets of simple objects.
2) Applying mean operation to each of the sets.
3) Recomposing a complex object from obtained simple objects.

The "effet Condorcet" (or paradox of voting) is another famous example of the difficulty met by such a procedure and the relation between these various aggregation paradoxes and the Arrow's theorem has been developed in the (unfortunately not sufficiently read) 1952 paper by Guilbaud."

Let me just add that in the present formulation I have been completely agnostic about the nature of the object to be called an aggregate. This radical decision makes it hard to give explicit meaning to attractive notions like that of a "mean", or that of "closeness". Otherwise, I feel very comfortable and happy to work within a model with such an old and distinguished tradition.

Within the realm of local and restricted aggregators, the theorem expresses a tension between two features which are essential in the description of a rule: these are its flexibility, and the way in which it distributes among agents the power to influence its result. We say that a rule is flexible if two conditions hold: one, that for each pair of alternatives, the rule's outcome is sometimes sensitive to changes in the relative positions of these two alternatives alone; and two, that this sensitivity should not be conditional on the position of any third

[1] Compare with my previous remark that "since not all combinations of characteristics for an aggregate are jointly compatible, the fully decentralized choices of characteristics for aggregators across different pairs of alternatives may be contradictory".

alternative. We say that one agent concentrates all power under a rule if only the preferences of this agent (when they are strict) determine the social aggregate.

I prove that if a preference aggregation rule is local and restricted, then either it is not flexible or else it gives all decision power to a single individual.

From this theorem we can derive as corollaries all the above mentioned classical results. Providing a common framework for such disparate theorems is interesting per se. In some cases (Arrow and Wilson), the close connection was obvious from the start. Yet, the present formulation allows for a better understanding of the role played by the Pareto condition in Arrow, and in particular reinforces Wilson's message that this role is quite limited. In the case of Arrow and Gibbard-Satterthwaite, the existence of a close connection has been very well understood for years, but claiming that both results are a corollary of a unique theorem is a stronger and novel statement. The connection established here between Sen's result and the rest may be more problematic because Sen's model and its interpretation are themselves subject to debate (Salles, 2000). At any rate, Saari (1998) had emphasized the relationship between Sen's and Arrow's result, by using a different line of argument. We discuss all these connections later, in Section 3.

I do not want to leave the reader with the impression that all I do here is to force old results into a common frame for the sake of it. I already pointed out that the theorem has content and meaning of its own: the tensions between locality, restriction and flexibility can only be solved by an extreme distribution of decision power. Moreover, this content is clearly reflected in the structure of its proof, which follows the essential line of two proofs of Arrow's and Gibbard-Satterthwaite's theorem that I provided some twenty years ago (Barberà, 1980, 1983a, 1983b). These separate proofs are now blended into a single one, which reveals the essence of the difficulty at hand. Local aggregation requires the decomposition of our choice of aggregate into partial choices, each of which must be exclusively dependent on the relative positions of some pair of alternatives. Although possibly centralized among agents, decisions are decentralized in terms of the influence exerted by different pairs of alternatives in determining which aggregates are ruled out as possible candidates. Since not all combinations of characteristics for an aggregate are jointly compatible, the fully decentralized choices of characteristics for aggregators across different pairs of alternatives may be contradictory. Some connections must be ensured between the choices made for some pairs and the choices made by some others. It turns out that, for any given preference profile, and as long as rules are flexible, this restricts the ability to influence the choice of aggregate to a single agent. Moreover, this local concentration of power extends to become global: only a single agent can make a difference, at any preference profile (as long as this agent holds strict preferences). The need for coordinated action becomes extreme: only one agent can make any difference at all!

Before I finish this Introduction, let me comment on some related papers. The connections between Arrow's theorem and the issue of manipulability had been remarked (see Vickrey, 1960) even before Gibbard's formulation of his celebrated theorem. Gibbard's proof relies on Arrow's, and establishes a connection that was later stressed by Satterthwaite (1975), Mueller and Satterthwaite (1977) and many other authors. One should point out that the possibility of defining a nondictatorial Arrowian Social Welfare Function on a given domain of preferences, or that of defining a nondictatorial strategy-proof social choice function on the same domain are logically independent questions: there are domains where one of the two objectives can be achieved and not the other. But the connection is tight for universal domains, the ones considered here, and also for the domains of single-peaked preferences (Barberà, Gul and Stacchetti, 1993).

For universal domains, the present theorem offers a convincing explanation of why so many parallels have been observed between the two classics: there is a general statement that covers both!

Regarding proofs, Arrow's own (Arrow 1963), which has been refined by many other authors (see Sen, 1970, for a very nice version), relies on a careful analysis of the global decision power of coalitions: it is first observed that, under Arrow's conditions, if a coalition is decisive for a pair it must be decisive for all pairs; it is then proven that if a coalition is decisive and can be broken into subcoalitions, one of these subcoalitions must also be decisive; finally, observing that the grand coalition is decisive allows to prove the theorem. In the preceding argument, decisiveness is the ability to determine the outcome of aggregation at all preference profiles: a coalition is either decisive or not. As already mentioned, Gibbard's (1973) early proof, and many that followed, used a construct that allowed to prove that result by indirectly appealing to Arrow's theorem. There exist many proofs of the Gibbard-Satterthwaite theorem, using different approaches. Those by Schmeidler and Sonnenschein (1978), Batteau, Blin and Monjardet (1981) or Barberà and Peleg (1990) did constitute real innovations by the time they appeared. Recent interest on the subject is obvious: three simple proofs of the theorem have been published recently (Benoit, 2000; Sen, 2001; and Reny, 2001). One of them (Reny, 2001) highlights the parallelism between the proofs of Arrow and (one of the) proofs of Gibbard-Satterthwaite. A very recent paper by Eliaz (2001) makes the point that both theorems are corollaries of a larger result. In that sense, the paper by Eliaz emphasizes, as my paper does, the fact that there is a deeper connection between these results than just a parallelism. Other than that, however, the paper by Eliaz and the present paper proceed quite differently, both in the formal model and in the type of proof.

My strategy of proof here follows the line of my already mentioned papers in the early eighties (Barberà 1980, 1983a, 1983b), which were the first to exploit the role of pivotal voters. The proof starts by considering the ability of agents to change the outcomes of the aggregation process at a given preference profile: it concentrates on a local property. Then I proceed to show that, under the conditions of the paper, if one agent can affect the outcome of the aggregation process at some profile, it must be able to affect it at any profile.

The paper is organized as follows. First, I present notation, formal definitions and a statement of the theorem in a simplified framework (Section 2). Then I discuss the relationship between the general theorem and the classical results (Section 3). Section 4 concludes, with a discussion of the possibility for a more general version of the theorem and its proof.

2. A simple version of the Theorem

Let $N = \{1,2,...,n\}$ be a finite set of *agents* ($n \geq 2$).
Let $X = \{x, y,...\}$ be a finite set of *alternatives* ($|X| \geq 3$).
Let $A = \{a_1, a_2,...\}$ be a set of *aggregates*.

2^A stands for the power set of A. $\mathcal{A}$ is the set of singletons in 2^A.
$\mathcal{D}$ stands for the set of distinct pairs of alternatives.

Let $\mathcal{P}$ be the set of linear orders[2] (complete, antisymmetric, transitive binary relations) on X. Elements of $\mathcal{P}$ are called *preferences*.

Elements of $\mathcal{P}^n$ are *preference profiles*. Preference profiles are denoted by P, P', etc. Given a preference profile P, and any set S of alternatives, denote by $P(S)$ the profile of preferences on S such that, for all $i \in N$, $P_i(S)$ is the restriction of i's preference P_i to the set S. In particular, we denote by $P(xy)$ the preference profile on the pair x, y induced by the profile P on X.

Alternatives x and y are *contiguous* in P_i iff for all $z \in X$, $z \notin \{x, y\}$, $zP_ix \leftrightarrow zP_iy$.

Alternatives x and y are contiguous in P iff they are contiguous in P_i, for all $i \in N$. Profiles P where two alternatives are contiguous play an important role in my setup, because they can be changed to other profiles P' where $P(xy) \neq P'(xy)$ while keeping $P(zw) = P'(zw)$ for all other pairs z, w other than x, y.

Profiles P and P' are *xy-equivalent* iff $P(xy) = P'(xy)$. Otherwise, say that P and P' differ on x, y. Profiles P and P' *only differ on x,y* if they differ on x, y, and they are zw-equivalent for any pair z, w other than x, y. Clearly, if P and P' only differ on x, y it must be that x and y are contiguous in the preferences of all agents, both in P and P'.

Profiles P and P' *differ on i* iff $P_i \neq P_i'$. They *only differ on i* iff $P_i \neq P_i'$ and $P_j = P_j'$ for all $j \neq i$. They *only differ on x,y for i,* if they only differ on i and they only differ on x, y.

Contiguity of a pair of alternatives is a special case of what I will call connectedness of a set of alternatives with respect to a preference. Formally, $B \subset X$ *is connected in* P_i if only if for all $x, y \in B$, $z \in A \setminus B$, $zP_ix \leftrightarrow zP_iy$. B is connected in P if it is connected for all P_i's.

[2] Notice that I assume preferences to be strict (no indifferences are allowed among different alternatives, by antisymmetry). This simplifies the exposition.

An *aggregation rule* (on the universal domain $\mathcal{P}^n$) is a function

$$f : \mathcal{P}^n \to \mathcal{A}$$

assigning a single aggregate to each preference profile[3].

An aggregation rule f is *local* iff for each pair $\{x, y\} \in \mathcal{D}$, there exist functions g_f^{xy} : $\mathcal{P}^n \to 2^A \setminus \varnothing$, such that

$$f(P) = \bigcap_{\langle x,y\rangle \in \mathcal{D},} g_f^{xy}(P)$$

and

(1) for each $x, y \in X$, P, P',

$$[P(xy) = P'(xy)] \to [g_f^{xy}(P) = g_f^{xy}(P')],$$

that is, g_f^{xy} is a function of x and y's relative position in a profile, for any profile.

(2) the range r_g^{xy} of g_f^{xy} consists of two distinct sets of aggregates T_1^{xy}, $T_2^{xy} \subset A$, for all $x, y \in X$[4].

Locality is a demanding requirement on aggregation rules. Locality requires that the global decision on which aggregator to attach to any given preference profile should be decomposable into partial decisions, one for each pair of alternatives, depending only on the ranking of this pair of alternatives in the profile[5]. Each partial decision attached to one pair of alternatives is a set of aggregators, and the global decision is the unique aggregator belonging to all of the sets selected, for each pair of alternatives, given the preference profile. Locality

---

[3] Notice that I model the function as choosing a singleton set in $\mathcal{A}$, rather than an alternative in A. This is to keep as formal as possible.

[4] This second requirement could be relaxed to cover several additional interesting cases. In particular, I could allow for the image of some g_f^{xy}'s to consist of a single value. This would leave room to include the constant functions (and others) among the local we could also allow for more than two sets in the range. This would extend the reach of the theorem's applications to cover, among others, the case of Arrowian aggregation when social preferences can express indifference between pairs of alternatives. And it would require a careful modification of the way to define restricted aggregators (see my concluding remarks). I have chosen to keep this somewhat restrictive requirement (2) for expositional clarity.

[5] Of course, this is extremely close to the idea that the aggregation rule satisfies Arrow's Independence of Irrelevant Alternatives. But I prefer to refer to it by another term because (1) it applies to a larger context, and (2) it implies the description of aggregation as the intersection of sets, which is neither necessary nor implied by Arrow's formulation.

also requires that the set of aggregators $g_f^{xy}(P)$ associated to each xy-pair at profile P should only depend on the relative ranking of x and y in the preferences of the agents, at that profile. Finally, the range of the partial decision functions g_f^{xy} is restricted to a limited number of values. In the present formulation, the ranges of these functions are restricted to be dichotomic.

The rules g_f^{xy} associated to a local aggregation rule f are called f's partial selection rules. For any P, $g_f^{xy}(P)$ is either T_1^{xy} or T_2^{xy}. I denote by $\left[g_f^{xy}(P)\right]^{comp}$ the element in the range of g_f^{xy} which is *not* chosen at P.

Before we proceed, let us see how this formulation applies to the frameworks of other classical results.

Consider the simple version of Arrow's theorem where the preferences of agents, as well as the preferences for society (the aggregates) are required to be linear orders. First notice that under the assumption of Independence of Irrelevant Alternatives, the social order between any two alternatives x, y, at any preference profile P, will be determined only by the ranking of x and y in the preferences of individuals at P. Hence, given a preference profile and two alternatives, x, y, one of the two possibilities "x preferred to y" or "y preferred to x" will be excluded. The set of all orders of alternatives ranking x and y in the nonexcluded position will be the image of the function $g_f^{xy}(P)$ for each P. Indeed, the image $f(P)$ will be the order respecting the nonexcluded positions for all pairs. Hence, it is clear that any IIA social welfare function can be written as a local aggregation function, where the aggregate is the intersection of sets of nonexcluded orders; a similar reasoning applies in the framework considered by Wilson.

The analysis of strategy proof rules as local aggregators builds on two essential remarks already made by Gibbard (1973). One is that, when all agents agree that all alternatives in a set S are preferred by all agents to all alternatives in $X \setminus S$, then a strategy-proof social choice function must choose from S. (I assume here that the set of alternatives equals the range of a strategy-proof social choice function. If not, drop the alternatives not in the range and start again, with no loss of generality). This implies, in particular, unanimity: an alternative must be chosen if it is unanimously considered to be the best by all agents. The second remark is that, if an alternative x is chosen when x and y are the best two alternatives for all agents, then y cannot be chosen at any preference profile where the order of these two alternatives is preserved. Building on these two remarks, notice that the outcome at each profile can be written in the form demanded by a local aggregate. For each given profile P, and each pair x, y, determine which one of the two alternatives is definitely a non-candidate to be the aggregate at this profile. For this, check which one of the two alternatives would not be elected at the profile $P(xy)$ obtained from P and where, ceteris paribus, x and y become the top two alternatives from all agents.

Now, any strategy-proof social choice function can be written as a local aggregation function, where for each pair of alternatives x, y, either the set $X \setminus x$ or the set $X \setminus y$ are retained (the one not containing the alternative which is eliminated by pairwise comparisons). The aggregate corresponding to each profile must clearly be expressed as the intersection of all the sets of non-excluded alternatives, for each pairwise comparison.

Having illustrated the basic features of local aggregation rules, let us now go back to study some further requirements.

Given a local aggregation rule f, and a set of pairs $D \subset \mathcal{D}$ define

$$G(P, D) = \bigcap_{\langle x,y \rangle \in D} g_f^{xy}(P).$$

We can interpret $G(P, D)$ as the set of aggregates which are not excluded on the basis of the comparisons of pairs in D. Clearly, with this notation, $G(P, \mathcal{D}) = f(P)$, for all P.

We say that f is *xy-sensitive* at P iff

$$f(P) = G(P, \mathcal{D}) \neq G(P, \mathcal{D} \setminus \{x, y\}) \cap \left[g_f^{xy}(P) \right]^{comp}.$$

Thus, f is *xy*-sensitive at P if a change in the value of the *xy*-partial selection would change the value of the global selection of aggregate.

The notion of sensitivity is introduced because it is not always the case that the change in one of the partial decisions associated to one pair of alternatives can change the global outcome. This depends on the type of aggregate we consider. Since we want to cover different types of aggregates, and provide a result that is independent of the type of aggregate under consideration, we cannot be more specific on the connections between changes in the partial decisions and eventual changes in the aggregate. In Arrow's framework, where each partial decision determines the ranking of a pair of alternatives within the overall ranking of alternatives, any change in one partial decision is translated into a change in the aggregate. Yet, in Gibbard's framework, the same is not true. Starting from some preference profile, you may change the partial recommendation on a pair *x*, *y*, and yet see a third alternative *z* be elected before and after the change: the aggregation rule would not then be *xy*-sensitive at that profile.

We say that f is *restricted* iff, for all $x, y, z \in X$, and all profiles P where f is *xy*-sensitive and *yz*-sensitive, there exist sets T_r^{xy} and T_s^{yz} in the range of g_f^{xy} and g_f^{yz} for which

$$G(P, \mathcal{D} \setminus \{x, y\}, \{y, z\}) \cap T_r^{xy} \cap T_s^{yz} \notin \mathcal{A}.$$

To better understand this assumption, let us again go back to examples. In Arrow's framework, for each function g_f^{xy}, the range consists of two possible outcomes, each of which is a set of orders: the orders with x over y, on the one hand, and the orders with y over x, on the other. The fact that aggregates must be orders implies that our aggregation function must be restricted: indeed, take any profile and any triple x, y, z. If at the initial profile P we have x over y, then the set of orders with x over y appears in the intersection of values defining the aggregate at P. Then, the set of orders with y over z, and the set of orders with z over x, both in the ranges of g_f^{yz} and g_f^{xz}, respectively, will have an empty intersection with the rest of sets of orders defining the value of f at P. Likewise, if at P we had that y stood over x, then the choices of sets of orders with x over z and z over y would have led to the empty intersection required by restriction.

Thus, the condition called restriction in our general theorem corresponds in Arrow's theorem to the requirement of transitivity of the social aggregate.

The social choice functions in Gibbard-Satterthwaite are also restricted local aggregators: indeed, for each pair of alternatives, the ranges of the functions g_f^{xy} consist of two well-defined sets: all alternatives except x and all alternatives except y. Notice that, for any triple x, y, z of alternatives, and any profile P where f is xy-sensitive and also yz-sensitive, the intersection of all the partial selections at P, other than those for xy and yz, must contain x, y and z. But then, the choice of $X \setminus y$ from the range of g_f^{xy}, coupled with the choice of $X \setminus y$, as well, from the range of g_f^{yz}, would leave us with both x and z in the overall intersection. Since only singletons, not two-element sets of aggregates, can be in the range of f, we have identified a restriction to be satisfied by the social choice functions considered in the Gibbard-Satterthwaite setup.

Hence, a rule is restricted if not all partial choices for different pairs are compatible with an acceptable global choice, when these partial choices do make a difference. A restricted local aggregation rule is thus one for which the choices of the different sets of aggregates associated to each pair of alternatives is not completely free: some combinations of potential choices would not be able to produce any aggregate at all. Hence, some coordination among partial outcomes will be necessary in order to obtain one and only one aggregator for each particular profile.

We say that f is *flexible* iff for some triple of alternatives $x, y, z \in X$, there is some profile P at which

(1) The set $\{x, y, z\}$ is connected in the preferences of all agents at P,
(2) f is xy-sensitive and also yz-sensitive at P,
(3) $G(P, \mathcal{D} \setminus \{x, y\}, \{x, z\}) \cap (g_f^{xy}(P)) \cap (g_f^{xz}(P))^{comp} \neq G(P, \mathcal{D} \setminus \{x, y\}, \{x, z\}) \cap (g_f^{xy}(P))^{comp} \cap (g_f^{xz}(P))^{comp}$,

and

$$G(P, \mathcal{D} \setminus \{y,z\}, \{x,z\}) \cap (g_f^{yz}(P)) \cap (g_f^{xz}(P))^{comp} \neq G(P, \mathcal{D} \setminus \{x,y\}, \{x,z\}) \cap (g_f^{yz}(P))^{comp} \cap (g_f^{xz}(P))^{comp},$$

In words, condition 3 requires that f's *xy*-sensitivity and *yz*-sensitivity at P would be preserved by changes in the value of g_f^{xz}.

Flexibility encompasses two requirements. One is that, for each pair of alternatives, their relative position alone can sometime make a difference. The other, complementary requirement, is that if the relative positions of two pairs make a difference at the same profile, then the relative position of any third alternative does not affect the rule's sensitivity with regard to any of these two pairs.

In Arrow's framework, flexibility is an implication of the Pareto axiom and IIA. In Gibbard's and Satterthwaite's, the connection requires more care and it is discussed below, in section 3.4.

Agent i *concentrates all power under rule* f iff, for all preference profiles P, P' and all pairs x, y,

$$\left[P_i(xy) = P_i'(xy)\right] \rightarrow g_f^{xy}(P) = g_f^{xy}(P').$$

Notice that, then, we can write each partial function as $g_f^{xy}(P_i)$, and that then f can be written as $f(P_i)$, for all P.

Also notice that a dictator (with strict preferences) concentrates all power, but that an anti-dictator also does, and so do any agents whose preferences are the only relevant inputs to determine the outcome of an aggregation rule.

<u>Theorem</u>

If an aggregation rule is local, flexible and restricted, then there is a single agent i who concentrates all the power under f.

<u>Proof</u>

Let f be local, flexible and restricted. We first provide some useful definitions.

We say that agent i is *xy-pivotal* at preference profile P iff $\exists$ P' such that P and P' only differ in xy for agent i and such that $f(P) \neq f(P')$. (Clearly, if agent i is *xy*-pivotal at P, he is also *xy*-pivotal at P'.)

Remark

Remark that if P and P' only differ in xy for agent i, and f is xy-sensitive at P (and P'), then i is xy-pivotal at P (and P').

The proof is organized through a sequence of simple Lemmas.

Lemma 1

If i is xy-pivotal at profile P, P' only differs from P in j's ranking of $\{z,w\} \neq \{x,y\}$ and $f(P) = f(P')$, then i is still xy-pivotal at P'.

Proof

Just notice that, since $f(P) = f(P')$ and i is xy-pivotal , it must be that $G(P, \mathcal{D} \setminus \{x,y\}) = G(P', \mathcal{D} \setminus \{x,y\})$.

Lemma 2

There cannot be any profile P at which two different agents i and j are xy-pivotal and yz-pivotal, respectively, with x and y contiguous in P_i and y and z contiguous in P_j.

Proof

Suppose there was, and that $\hat{P}_i$, $\hat{P}_j$ were the preferences (with x and y contiguous in $\hat{P}_i$, and y and z contiguous in $\hat{P}_j$) such that $f(P_{-i}, \hat{P}_i) \neq f(P)$, $f(P_{-j}, \hat{P}_j) \neq f(P)$. Then, we can write the images $f(P)$, $f(P_{-i}, \hat{P}_i)$, $f(P_{-j}, \hat{P}_j)$, $f(P_{-ij}, \hat{P}_i, \hat{P}_j)$ as

$$G(P, \mathcal{D} \setminus \{x,y\}, \{y,z\}) \cap T_r^{xy} \cap T_s^{yz},$$

for all possible choices of T_r^{xy} in the range of g_f^{xy} and T_s^{yz} in the range of g_f^{yz} (use Lemma 1). Hence, one of these four possible outcomes cannot be an aggregate, because f is restricted.

Lemma 3

There is an agent who concentrates all power under rule f.

Proof

Since f is flexible, $\exists P$ and $x, y, z \in X$ such that
(1) the set $\{x, y, z\}$ is connected in the preferences of all agents at P, and
(2) f is xy-sensitive and also yz-sensitive at P.

Notice that, since f is xy-sensitive at P there must exist P', P'' such that they only differ on $\{x, y\}$ in i and $g_f^{xy}(P') \neq g_f^{xy}(P'')$. Similarly, there must exist $\widetilde{P}, \hat{P}$ and j such that they only differ on $\{y, z\}$ in j, and $g_f^{yz}(\hat{P}) \neq g_f^{yz}(\widetilde{P})$.

Since x, y and z are connected, it is possible to find a profile $\overline{\overline{P}}$ where x, y and z are still connected, where $\{x, y\}$ are contiguous in $\overline{\overline{P}}_i$, $\{y, z\}$ are contiguous in $\overline{\overline{P}}_j$, and i is xy-pivotal, while j is yz-pivotal. We use here condition 3 in the definition of flexibility, since in such profile the relative position of xz or zy may be different that at the original P. Only one of these two needs to be changed in order to meet the requirements. To see that, notice that if an agent is xy-pivot when x and y are contiguous, the agent is still pivot when its preferences over x and y are switched.

But now we have the conditions of Lemma 1. Hence, only one agent can be pivotal at this profile, and it is pivotal for two different pairs.

From this point on, we can start switching the preferences of any agent for different pairs of contiguous alternatives including any one of the three alternatives x, y or z in question. No such change alter the outcome if it happens for any agent other than i, since that would lead us to the situation of Lemma 2. Hence, at some profile we shall eventually get a pivot for the pair, and this pivot will necessarily be i again. That general argument can be applied repeatedly to conclude that the same agent i must be pivot at some profile for all pairs of alternatives and hence concentrate all the power under f.

3. Some applications and corollaries

In what follows, I present a sketchy discussion of why the new theorem does indeed cover the setups for which different results were initially formulated, and how the particular conditions under which the classical theorems hold are particular versions of those imposed here.

3.1. Arrow's framework and Theorem

We have already discussed why Arrowian Social Welfare functions are local and restricted. Notice now that the Pareto condition imposed by Arrow implies flexibility. Indeed, for profiles where all agents agree that x is preferred to y, x must be preferred to y socially, and the reverse must hold when all agents rank y over x. By changing the preferences one by one, reversing the order of x and y, starting from a profile where x and y are contiguous and all agents prefer x to y, and moving to one where they all prefer y to x, it must be that one agent is xy-pivotal at some of the profiles in this sequence. This can be done for other pairs yz. Moreover, xy-pivotality and yz-pivotality can be preserved in profiles where the conditions of flexibility do hold.

It is thus obvious that an Arrowian social welfare function must be a local, restricted and flexible aggregator: hence, all the decision power must be concentrated in an agent. Moreover, again by Pareto, this agent obtains the ranking x over y whenever he prefers x to y: this is the dictator.

3.2. Wilson's theorem

The initial framework is the same as Arrow's: IIA implies locality and transitivity implies that the aggregation function is restricted.

Unlike in Arrow's formulation, flexibility does not come from Pareto, but from the range condition requiring that x should be ranked above y for some profile, and below y for some other. The same construction as before, but having as extremes two such profiles with x and y contiguous, shows the existence of xy-pivots for all pairs. Hence, one agent always concentrates all decision power: but this agent may be getting always the same order among pairs of alternatives that he expresses, or the reverse. Because Pareto does not necessarily apply, the individual who concentrates the power may be a dictator or an anti-dictator. Other combinations, where the agent might be a dictator over some pairs, and an anti-dictator over other pairs are excluded by restriction. For example, if some pairs were ordered according to the preferences of one agent and others in the opposite direction than the preferences of the same agent, this could easily lead to cycles.

On the other hand, we may consider the constant function as an example of an aggregation rule which satisfies all of our conditions except for flexibility and locality[6]. This is the other type of aggregator obtained by Wilson, again violating the Pareto condition.

[6] If we had chosen to use a more general definition of locality as discussed in Aizerman and Aleskerov (1995), then the constant function could be accommodated among the local ones, while still violating flexibility.

3.3. Sen's impossibility of a Paretian liberal

The essential remark regarding Sen's result is that, it postulates explicitly as desiderata the two conditions which are stated by our theorem to be incompatible under local and restricted rules. One of them is the Pareto condition: we have already seen that this condition, for a universal domain of preferences, implies that the aggregation function must be flexible. The other condition imposed by Sen is that at least two agents must be able to determine the strict ranking of two alternatives each (these two alternatives describe social states which are identical except for some features over which the agent in question is entitled to have full control. This is the condition of (minimal) liberalism, a condition which is in open contradiction with the possibility of one agent holding all the decision power under an aggregation rule. The interpretation of this ability to determine the ranking of some alternatives is subject to debate, because Sen is referring to rights, and their connection to choices is not an obvious one. In the present paper I limit myself to comments about the formal structure in Sen's formulation, not about its meaning(s).

Since Sen requires two conditions which are incompatible for any local and restricted rule, it is enough for us to argue that he is considering, in his theorem, rules that can be written as members of this class. No explicit argument is provided by Sen regarding the general structure of the functions he considers: hence, some characteristics of these functions for some special profiles might escape the description (and hence the full force) of our theorem. But remember that the theorem contains two statements: one is about the impossibility of allocating pivotal rights to more than one individual at the same profile (the local aspect, step one of the proof), and the other is the extension of this local argument to a global one. We only need the local result to understand Sen's impossibility. When it applies, the Pareto principle allows us to ascertain that the aggregate must be among the class of binary relations respecting the agents' unanimous ranking of a pair. Likewise, the rankings of pairs associated to the private domain of an agent depend only (by the assumption of minimal liberalism) on the interested agent's ranking of these pairs. Hence, the aggregate for profiles where Pareto applies and agents have strict preferences over the pairs of their exclusive concern must be in the intersection of those sets of aggregates respecting the above conditions, as in a local aggregation rule. Now, the aggregate is required by Sen to be an acyclic rule, and this imposes a restriction on the choice of aggregates, which cannot be respected at profiles where Pareto applies and both agents with rights can exert them.

3.4. The Gibbard-Satterthwaite theorem

I have already discussed why a strategy-proof social choice function is a local and restricted aggregator. Moreover, strategy-proofness guarantees that, whenever the range of the rule contains at least three alternatives, then there must be profiles where some agent is pivotal for more than one pair. My 1983 proof of the Gibbard-Satterthwaite theorem establishes (step g) the following consequence of strategy-proofness: if no individual is ever pivotal under f for more than two alternatives, then the range of f consists of at most two alternatives. Moreover, the construction in Gibbard's proof that determines whether each alternative is excluded or is not at a profile can only depend on the positions of pairs of alternatives, when f is strategy-proof. This gives us the second part of flexibility. We can now refer to the general theorem to obtain dictatorship.

4. Final Remarks

I have proven the theorem for the case where the range of each partial selection rule consists of two sets only. The proof is sufficient to cover the Gibbard-Satterthwaite theorem. It is also sufficient to cover the versions of the rest of the theorems we mention (Arrow, Wilson, Sen) where society's preferences are required to be strict. But since these theorems also hold when societies (as well as individuals) are allowed to express indifferences, it would be interesting to discuss how our theorem should be extended to cover the case of three-valued (or many-valued) ranges. The same basic ideas apply, but further definitions and some delicate distinctions must be made. I leave this extension for another paper.

REFERENCES

Aizerman, M. and F. Aleskerov (1995): "Theory of Choice", North Holland, Elsevier Science, Amsterdam.

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

Barberà, S. (1980): "Pivotal Voters: A New Proof of Arrow's Theorem", *Economics Letters* 6, 13-16.

Barberà, S. (1983a): "Strategy-Proofness and Pivotal Voters: A Direct Proof of the Gibbard-Satterthwaite Theorem", *International Economic Review* 24, 2, 413-418.

Barberà, S. (1983b): "Pivotal Voters: A Simple Proof of Arrow's Theorem", in P.K. Pattanaik and M. Salles (Eds.), *Social Choice and Welfare*, North Holland.

Barberà, S., Gul, F. and E. Stacchetti (1993): "Generalized Median Voter Schemes and Committees", *Journal of Economic Theory* 61, 262-289.

Barberà, S. and Peleg, B. (1990): "Strategy-proof Voting Schemes with Continuous Preferences", *Social Choice and Welfare* 7, 31-38.

Batteau, P., Blin, J-M. and B. Monjardet (1981): "Stability of Aggregation Procedures, Ultrafilters and Simple Games", *Econometrica* 40, 527-534.

Benoit, J-P. (2000): "The Gibbard-Satterthwaite Theorem: A Simple Proof", *Economics Letters* 69, 319-322.

Eliaz, K. (2001): "Arrow's Theorem and the Gibbard-Satterthwaite Theorem Are Special Cases of a Single Theorem", mimeo, Tel-Aviv University.

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result", *Econometrica* 41, 587-601.

Guilbaud, G. Th. (1952): “Les Théories de l’Intérêt Général et le Problème Logique de l’Aggrégation“, *Economie Appliquée* 5, 501-584. English translation in P.F Lazarsfeld and N.W. Henry (Eds.) *Readings in Mathematical Social Sciences* (Science Research association Inc.-Chicago, 1965) pp. 262-307.

Mueller, E. and A. Satterthwaite (1977): "The Equivalence of Strong Positive Association and Strategy-Proofness", *Journal of Economic Theory* 14, 412-418.

Reny, P.J. (2001): "Arrow's Theorem and the Gibbard-Satterthwaite Theorem: A Unified Approach", *Economics Letters* 70, 99-105.

Saari, D. (1998): “Connecting and Resolving Sen’s and Arrow’s Theorems”, *Social Choice and Welfare* 15, 239-261.

Salles, M. (2001): "Amartya Sen: Droits et Choix Social", *Revue Economique* 81-3, 445-457.

Satterthwaite, M. (1975): "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", *Journal of Economic Theory* 10, 187-217.

Schmeidler, D. and H. Sonnenschein (1978): "Two Proofs of the Gibbard-Satterthwaite Theorem on the Possibility of Strategy-Proof Social Choice Functions", in H.W. Gottinger and W. Leinfellner (eds.), *Decision Theory and Social Ethics, Issues in Social Choice*, 227-234, D. Reidel Publishing Company, Dordrecht, Holland.

Sen, A.K. (1970): "The Impossibility of a Paretian Liberal", *Journal of Political Economy* 78, 152-157.

Sen, A. (2001): "Another Direct Proof of the Gibbard-Satterthwaite Theorem", *Economics Letters* 70, 381-385 (2001).

Vickrey, W.S. (1969): "Utility, Strategy and Social Decision Rules", *Quarterly Journal of Economics* 74, 507-535.

Wilson, R. (1972): "Social Choice without the Pareto Principle", *Journal of Economic Theory* 5, 478-486.