Social Choice Theory

Jean-Pierre Benoît*

1 Prologue

In June of 1987, following weeks of rioting, the ruling party in South Korea announced constitutional changes which included provisions for the next president to be chosen through direct elections. Anti-government feelings were high and ruling party candidate Roh Tae Woo was expected to be defeated easily. Nonetheless, Roh emerged victoriously in the election that took place later that year. The final vote totals for the major candidates were:

36.5%	Roh Tae Woo
27%	Kim Dae Jung
26%	Kim Jong Pil

The election was marred by allegations of improprieties, but the real reason for Roh's victory appears to have been the split among the two Kims. Although the vote totals only tell us about each voter's most preferred candidate, it seems likely that a more complete listing of preferences was something like:

36.5%	27%	26%
Roh	Kim D.J.	Kim J.P
?	Kim J.P	Kim D.J.
?	Roh	Roh

Note that with these preferences, either of the losing Kims could have beaten the winner Roh in a one-on-one election, with 53% of the vote. Although the president was chosen by a direct vote, the winner does not seem to have been truly representative of the people. Perhaps this was due to a flaw in the voting system that was used.

2 Introduction

Who should be mayor of New York City? Should the workers at a certain textile plant be represented by union A, union B, or no union at all? How much should a sick person have to pay for his medical care? In which restaurant should a group of friends dine? Who was the most valuable baseball player in the National League last year? How wide should lapels be next year? Which drugs should be legal to consume?

The above questions differ in nature and importance, but from these differences two important similarities emerge. First of all, in each case people are likely to disagree as to the proper answer. Second of all, answers are somehow reached. The methods by which these answers are reached vary widely: various voting systems, economic markets, implicit conventions.... but are these methods "good" methods, and what would we mean by a good method, anyway?

The basic query of social choice theory is the following: How should we combine the preferences of the various members of society into a "preference" for society as a whole? New York City residents may have different candidates they would like to see as their next mayor, but ultimately just one person, who presumably is to represent "the will of the people", shall be chosen. Is it possible to make this notion of the will of the people precise? Would we get a "better" representative by selecting the top vote-getter in a multi-candidate election, even if he or she does not have a clear majority, or by cutting down the number of candidates until someone emerges with a majority?

To address these questions we take an indirect approach. We consider what properties we might desire any method of choosing an alternative to possess. For instance, one obvious property would be that if it just so happened that everybody agreed on which alternative was the best, then that alternative should be selected. Just about every voting procedure in use satisfies this property. However, we will see that many seemingly obvious properties turn out not to be possessed by common voting systems. In fact, several apparently desirable properties are actually in conflict with each other.

We all have some familiarity with the peculiarities of politics and the possible anomalies that can occur in group decision making. Our findings will indicate that many of these peculiarities and seeming anomalies are actually unavoidable.

3 Arrow's Impossibility Theorem

3.1 Social Decision Rules

To begin our analysis, it is useful to recast our questions in a more general framework. We start with a society composed of distinct individuals (e.g., the residents of New York City). They must make a choice from among certain alternatives (e.g., five mayoral candidates). Each individual has a ranking of these alternatives from most preferred to least preferred (e.g., John Jefferson's most preferred candidate may be the Democrat incumbent, followed by one Democrat challenger, followed by another Democrat challenger, followed by the sole Republican candidate and the lone Socialist; Mary Ortiz may favor the Republican, followed by the Democrat incumbent, followed indifferently by both Democrat challengers, and lastly the Socialist; Betty Jones may favor ... and so forth).

The problem of Social Choice Theory is how to choose an outcome for society as a whole from the rankings of the individual members of society. That is, given that people are likely to disagree, how should we combine, or aggregate, their individual preferences into a single choice which tells us what society prefers. Notice that all the questions we first considered can be put into this framework. For instance, each dining friend has preferences over restaurants. One way of deciding upon a restaurant would be to first let each person veto any type of food and then choose the restaurant that is favored by the most people. Another way would be to simply go along with the preferences of the eldest diner.[1]

Now we could approach our social decision problem by just proposing some decision method straight off, say plurality rule or a plurality vote followed by a runoff. But there is a problem with such an approach. Although our proposed methods might seem relatively straightforward, they may, in fact, possess many unanticipated features.

For instance, consider one-on-one majority comparisons. Let us say a group of workers must decide between a militant union, a moderate union, and no union at all. Thirty workers are opposed to any union, but if they must have one, they prefer a moderate union to a militant one. Thirty four workers desire a union, favoring first a moderate union, then a militant union. Finally, thirty two workers believe that a moderate union would be the worst

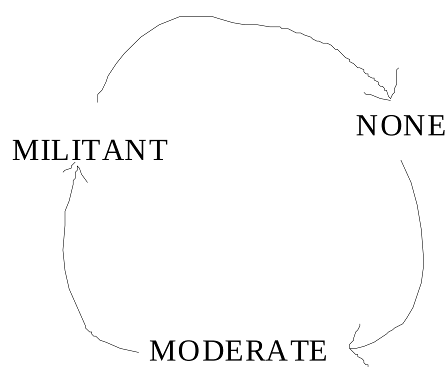
[1] In Quaker meetings, disagreements are resolved by "consensus". This mechanism does not fit easily into our framework.

possible option, being likely to draw the workers into strikes they will not have the mettle to win. These workers most favor a strong militant union, but then would like no union at all. We have the following rankings:

The Condorcet Voting Paradox

	30 workers	34 workers	32 workers
Best	None	Moderate	Militant
Second	Moderate	Militant	None
Worst	Militant	None	Moderate

The workers vote on the alternatives two at a time. First the moderate union beats the militant union, 64 to 32. Then, the no-union option beats the moderate union 62 to 34. The workers end up with no union. But is this really what they want? Notice something bizarre. If we put the militant union up against the no-union option, the militant union wins 66 to 33! But the militant union already lost out to the moderate union. Plurality rule leads to a cyclical ranking of the alternatives:



So which alternative is the best? Having no union is both better and worse than having a militant union. Should we declare all the options tied and pick one at random? For now, let us just note that starting directly with a decision rule begs the question of what properties the decision rule possesses. Instead, let us follow a more basic approach, and first decide what properties we would like our voting method to possess.

To make the problem interesting, we will always assume that there are at least two individuals in our society, and that they have at least three alternatives among which they must choose.[2] In addition, since we cannot anticipate individuals' preferences, we will insist that our decision rule be ready to give us a social ranking for any pattern of individual rankings.

We wish to decide what properties a good decision rule should possess. We begin with two simple properties.

- **Property 1**: **Unanimity**
  If everyone agrees that alternative A is the best alternative, then alternative A should be selected.

This property seems almost too obvious to be stated. If everyone agrees that Maradona was the greatest soccer player of all time, then this should also be society's official verdict. Note that this condition does not have much bite since it is rare to find complete agreement among voters.

- **Property 2**: **Non-Dictatorship**
  There does not exist a person – a *dictator* – such that society's choice is always that person's top choice, regardless of the preferences of others.

Although there may be (too) many among us who harbor dictatorial aspirations, a decision rule that reflects only one person's views can hardly be considered representative.

Now if an individual's favorite choice for mayor from a group of five candidates is the socialist entrant, naturally that individual also prefers the socialist to each of the other candidates taken one at a time. Similarly, if I consider Maradona to be the greatest soccer player ever, there is no point in saying, "yes, but let us compare Maradona only to Pelé, who is better?" My answer will remain Maradona. We might expect the same from a social decision rule. That is, we might want that whatever outcome our decision rule selects should also beat all the other outcomes in pair-wise comparisons. It would be distressing if after deciding that A was preferred to B, C, and D, we put A up against D alone and, without anyone's preferences changing,

[2] In Section 4 we consider the case of two alternatives.

we now decided that D was better than A. The claim that A was initially representative of people's desires would seem weak, given that these same people choose D over A when given the chance. This is what happened in the 1987 Korean election, where Roh was the winner although Kim D. (and Kim J. for that matter), could have beaten him in a one-on-one contest. Thus, we have:

- **Property 3: Pair-wiseness**
  The winning outcome(s) from a set of alternatives should also beat all the other alternatives in pair-wise contests.[3]

One way of thinking of pair-wiseness is that it requires that a winning candidate still be chosen if some of the other candidates drop out of the running. Thus, we could say that a pair-wise decision rule is "independent of the presence of losing candidates".

Although pair-wiseness is not as obviously a desiratum as unanimity and non-dictatorship, it is, to say the least, seemingly a reasonable condition. Indeed, it seems to be a not very demanding property and it would appear that we should impose many more conditions than the aforementioned three if we want to ensure that our decision rule is acceptable. For instance, in addition to the unanimity property we might want to require that if everyone but one person considers A to be the best alternative, then society should as well. Nevertheless, Kenneth Arrow proved that rather than needing to add more conditions, we have already gone too far. We state this remarkable result as theorem 1 below. First let us be a little more precise.

Our basic data are a set of *alternatives*, a *group of individuals*, and the individuals' *rankings* of these alternatives. We call a set of rankings, one for each individual, a *preference profile*. Given any preference profile, a *decision rule* selects a single alternative, or declares a tie, on the basis of these rankings. The decision rule operates on a subset of the alternatives, e.g. any pair, by selecting one of them on the basis of the rankings over this subset of alternatives. (Note, that we do not rule out the possibility that the rule operates "differently" on different sized subsets.)

Consider every possible preference profile. A decision rule is **dictatorial** if there exists one individual whose top choice is selected in every case. If for each individual there exists at least one preference profile under which her favorite choice is not chosen, then the decision rule is non-dictatorial. A

[3] Any two winning outcomes should be tied in their pair-wise contest.

decision rule is **unanimous** if for those profiles in which everyone agrees on the top alternative, that alternative is selected. A decision rule is **pair-wise** if for every preference profile the outcome(s) selected from the set of alternatives is also selected in all of its two-way contests. If there is a preference profile for which the selected outcome loses to another outcome in a pair-wise contest, then the rule is not pair-wise.

Note that if a decision rule is pair-wise, the winning alternative(s) can be determined on the basis of pair-wise comparisons alone.

Theorem 1 *(Arrow's Impossibility Theorem): Suppose there are at least three alternatives from which to choose and at least two individuals in society. Then there does not exist a non-dictatorial, unanimous, pair-wise decision rule.*
Alternatively, the only way to ensure that:
1) a unanimously preferred alternative is selected whenever such an alternative exists, and
2) the selected outcome would always be picked above the other choices in pair-wise contests
is to appoint a dictator.

In a slightly different approach we could begin by defining a rule directly on paired contests alone. The idea would be to declare the overall winner to be the alternative that wins all the two-way contests. This would give a well-defined rule if there were never any cycles (at least at the top). Arrow's theorem, however, implies that any rule for two-way contests will sometimes produce such cycles. This theorem, then, is "simply" a generalization of the Condorcet voting paradox.

3.1.1 Another Look

To further our understanding of Arrow's theorem, let us present it in a different formulation. We motivate this formulation with an example.

The 1980 United States presidential election saw the Republican candidate Ronald Reagan running against the Democratic incumbent Jimmy Carter and the independent candidate John Anderson. As Anderson campaigned, Reagan supporters looked on with glee. They reasoned that Anderson, a relatively liberal candidate, would draw votes away from Carter to Reagan's benefit.

This points out a peculiar aspect of the plurality election method. When a candidate's support increases this may result not in him or her being elected, but a third person altogether. In the 1980 election there was the possibility of Carter winning if Anderson's support was minimal, but Reagan winning if Anderson had strong support (thus splitting the more liberal vote). But why should we consider Reagan a more desirable candidate simply because someone else's popularity increases? (as it turned out Reagan did not need Anderson's help, as he won easily – but that's another story.)

Thus, we might ask the following of our decision rule. Suppose, that for some preference profile alternative A is selected. Now consider a different profile just like the original one, except that some people now rank some alternative X higher than they did before. We would like that this improvement in X's ranking not benefit anyone other than X. That is, the decision rule should now choose X, or still choose A if the increase in X's popularity has not been sufficient. In any case, this improvement in X's position should not result in some third alternative, say C, now being chosen.

Though this might seem like a reasonable property[4], theorem 2 indicates that it is too much to ask for of a *single-valued* decision rule (that is a decision rule that never declares any ties but, rather, always picks a single winner).

Theorem 2 *(Muller-Satterthwaite) Suppose there are at least three alternatives and at least two individuals. Then the only unanimous single-valued decision rule that ensures that an increase in one alternative's popularity can never work to another alternative's benefit, is a dictatorship.*

3.2 Social Welfare Functions

Arrows original formulation was different than the one presented above. Rather than looking for a decision rule which selects one (or more) winning alternative, he defined a **social welfare function**, which provides a complete ranking of all the alternatives for society based upon the rankings of the individual members of society. That is, given the individuals' rankings of the alternatives from first to last, a social welfare function also ranks them from first to last (allowing for the possibility of ties).

[4]This property, in contrast to pair-wiseness, does not require our decision rule to be defined for pair-wise comparisons as well as over the entire set of alternatives. Despite this difference, the two properties are essentially equivalent.

To each of the properties in theorem 1 for decision rules corresponds a property for social welfare functions. Thus, we have:

- **Unanimity**: If everyone ranks A above B, then society ranks A above B.
- **Non-Dictatorship**: There is no individual such that society's ranking is that individual's ranking, regardless of the preference profile.

Corresponding to our pair-wiseness condition is a condition known as *independence of irrelevant alternatives.* Pair-wiseness insists that the top choice also beat all other alternatives pair-wise. Thus, a determination that the top choice, say A, is better than one of the other choices, say B, can depend only on how people rank A compared to B, and not on how these options further compare to other "irrelevant" alternatives. This is because when we put A against B in a pair-wise contest, we consider only the individuals' rankings of these two alternatives. In terms of the social welfare function the condition is:

- **Independence of Irrelevant Alternatives**: Suppose that for some preference profile, the social welfare function ranks A above B. Now consider a second profile in which every individual's relative ranking of A and B is the same as in the first profile. The social welfare function must rank A above B for this second profile as well.

Theorem 1' (Arrow's Impossibility Theorem): *Suppose there are at least three alternatives from which to choose and at least two individuals in society. Then there does not exist a non-dictatorial, unanimous, social welfare function that is independent of irrelevant alternatives.*

3.3 Pair-wiseness and Independence Reconsidered

Unanimity and non-dictatorship are obviously highly desirable properties. The same cannot be said of pair-wiseness or independence of irrelevant alternatives.

Consider someone who is hosting a dinner with two guests and must decide whether to serve coffee, tea, or lemonade (he cannot afford more than one). Although the choice is only among coffee, tea, and lemonade he asks his two guests for a more complete ranking. He decides that if his guests indicate the following preferences:

Astrud	Gerald
Coffee	Tea
Tea	Juice
Beer	Rum
Juice	Cognac
Rum	Beer
Cognac	Coffee
Lemonade	Lemonade

he will serve tea. On the other hand, if his guests indicate the following rankings:

Astrud	Gerald
Coffee	Tea
Beer	Coffee
Rum	Cognac
Juice	Rum
Tea	Beer
Cognac	Juice
Lemonade	Lemonade

he will serve coffee. Notice that the host's welfare function rule violates independence of irrelevant alternatives. The ranking of the beverages coffee and tea is the same in the two profiles, namely, Astrud prefers coffee to tea, Gerald prefers tea to coffee, but the host's ranking of these beverages changes. By the same token, a pair-wise decision rule would require that the same beverage be served in both cases since the pair-wise comparisons are identical.

Nevertheless, the host's reasoning is clear. Firstly, lemonade is eliminated both times since it is unanimously dispreferred to both coffee and tea. The host then notes that in the first case, Astrud apparently likes both coffee and tea, while Gerald appears to like tea and not care too much for coffee, so that tea seems fair. In the second case, the reverse is true, so that coffee seems fair. The extra "irrelevant" alternatives appear to have provided additional information beyond the simple statements that Astrud prefers coffee to tea and Gerald prefers tea to coffee. But is this reasoning really valid? Consider the first case again. Perhaps the listings given correspond to the following preferences:

	Astrud	Gerald
Like	Coffee	Tea
		Juice
		Rum
		Cognac
		Beer
		Coffee
		Lemonade
Hate	Tea	
	Beer	
	Juice	
	Rum	
	Cognac	
	Lemonade	

so that actually coffee should be served. The point is that the extra information the host is getting might not be what he thinks it is.[5] Furthermore, why introduce these four irrelevant alternatives? What would have happened if a different four had been introduced, or a hundred instead of four? If irrelevant alternatives can count we are asking for trouble.

For another example, consider a hungry group of friends who decide that a Haitian restaurant is preferable to either a French restaurant, or a Chinese restaurant. On the way to the Haitian restaurant they discover that the town's only Chinese restaurant is closed, so they change plans and opt for French food. This violation of independence certainly seems a little peculiar, to say the least.

3.4 An Avoidable Peculiarity

Theorem 2 indicates that any unanimous, non-dictatorial decision rule has the feature that an increase in one alternative's popularity will sometimes benefit another alternative. For instance, consider an election with a liberal, a moderate, and a conservative candidate. The populace divides into three

[5] An economist would say that rankings provide only ordinal, not cardinal information, and that interpersonal comparisons cannot be made.

groups with the following preferences:

(A)

Group I	Group II	Group III
35%	28%	37%
Cons.	Mod.	Lib.
Mod.	Lib.	Mod.
Lib.	Cons.	Cons.

With a plurality rule, the liberal is elected. Now suppose that the moderate's ranking increases with some of group III's voters, so that the new profile is:

Group I	Group II	Group III	Group III'
35%	28%	33%	4%
Cons.	Mod.	Lib.	Mod.
Mod.	Lib.	Mod.	Lib.
Lib.	Cons.	Cons.	Cons.

The increase in the moderate's ranking works to the benefit of the conservative – although the ranking of this latter candidate has not changed, she is now elected. An even more bizarre occurrence would have been if an increase in the liberal candidate's own ranking had harmed her herself. That never could happen with the plurality method. Starting with the preference profile in (A), any increase in the liberal's ranking will only cause her to win more decisively. However, consider a plurality followed by a runoff (i.e., if no one wins a clear majority in the first round, the top two vote-getters are paired off in a runoff election).

We begin with the preference profile of (A). In the first round the liberal and the conservative candidates get the most votes (37% and 35%) and advance to the runoff. The liberal candidate wins the runoff with 65% of the vote.

Now suppose that before the first election the liberal actively engages in campaigning aimed specifically at those voters who dislike her most. She successfully persuades some group I members to now rank her first, so that the new population split looks as follows:

Group I'	Group I	Group II	Group III
8%	27%	28%	37%
Lib.	Cons.	Mod.	Lib.
Cons.	Mod.	Lib.	Mod.
Mod.	Lib.	Cons.	Cons.

Note that the only difference between this preference profile and the original preference profile is that some people now rank the liberal candidate higher than they did before. But what happens in this election? In the first round the liberal and the moderate candidate receive the most votes (45% and 28%). In the second round the moderate wins with 55%. The liberal candidate's increased popularity has actually harmed her! This possibility certainly gives candidates peculiar incentives

4 Majority Rule

Arrow's impossibility theorem assumes that society has at least three alternatives available. What about when there are only two alternatives? In this case there is no possibility of a cycle. Majority rule is then feasible and there does seem to be something especially "fair and democratic" about it. This aspect of majority rule can be made precise as follows.

A system is called anonymous if it does not favor any one person over another. In an anonymous system each person's preference counts equally. Examples of non-anonymous systems are corporate decisions where a large shareholder gets more votes than a small one, and voting in the United Nations general assembly, where permanent members of the security council have veto powers that the other countries do not have. In an anonymous system, if you interchange the preferences of two people, the outcome remains the same.

A system is called neutral if no particular outcome is favored over another. An example of a non-neutral system is one where a 2/3 majority is needed to overturn the status quo. In an anonymous system it does not matter how people are labeled, whereas in a neutral system it does not matter how alternatives are labeled. These two properties reflect the principles of one-person-one-vote and all-candidates-have-an-equal-chance.

Now a non-perverse system should reflect the preferences of individuals in a positive way. That is, first suppose society declares that A is at least as good as B. Now some individual who used to rank B above A, decides that she actually prefers A to B, whereas everyone else sticks to their previous judgment. Society should now declare that A is preferred to B (i.e. if previously A was considered better than B, it should remain better, whereas if it was just indifferent to B, this change should tip the balance.) This condition is known as positive responsiveness.

Under majority rule, alternative A is chosen over B if more people prefer A to B than the opposite. It turns out that the above three conditions – anonymity, neutrality, and positive responsiveness – completely characterize majority rule. That is, majority rule satisfies these conditions, and furthermore it is the only choice rule that does.

Theorem 3 *When deciding between two alternatives, majority rule is the only method that is anonymous and neutral, and satisfies positive responsiveness.*

The proof is relatively simple. From anonymity and neutrality, the ranking of alternative A relative to B can depend only on the number of people who prefer A to B (and not on the nature of A and B, or exactly who prefers A to B). Hence, if two alternatives receive the same number of votes, they must be tied. By positive responsiveness, then, if one of the alternatives receives more votes it must be preferred. But this is just majority rule.

Thus, we see that there are good reasons for the popularity of majority rule. Loosely speaking, if there are only two alternatives (and only one election), it is the only reasonable unbiased rule. When there are more than two choices, however, pair-wise majority rule may fail to produce an unambiguous winner, as the Condorcet voting paradox shows.

5 Condorcet Consistency

A **Condorcet winner** is an alternative that obtains a majority of the vote in every pair-wise contest against the other alternatives. As we have seen, a Condorcet winner need not exist. To the extent that majority rule is otherwise desirable, we might want to consider voting rules that yield the Condorcet winner whenever such a winner exists. A decision rule is called Condorcet consistent if it always yields the Condorcet winner when such a winner exists. Thus, different Condorcet consistent rules differ only in the selection they make when there is no Condorcet winner. Three Condorcet consistent rules are:

1. *Voting by Successive Elimination*
   The M alternatives are numbered from 1 to M. Alternative 1 is put up against alternative 2 in a majority rule contest. The winner is then put against alternative 3. The winner of this contest is then put against alternative 4 and so forth. The winner of the last contest is chosen.

Note that a Condorcet winner will win all the contests in which it is entered, and hence will be the winner of the last contest.

2 The *Copeland Rule*
Compare A with every other candidate x. Score $+1$ if a majority prefer A to x, -1 if a minority prefer A to x, and 0 if A and x are tied. A's Copeland score is the sum of these scores. The Copeland winner is the alternative with the highest Copeland score.

Note that a Condorcet winner will be a Copeland winner with a score of $M - 1$.

3 The *Simpson Rule*
Consider A, and for every other candidate x, compute the number $n(A, x)$ of voters who prefer A to x. A's Simpson score is the smallest value of $n(A, x)$, considering all x's. The Simpson winner is the alternative with the highest Simpson score.

Note that with N voters, a Condorcet winner will be a Simpson winner with a score greater than N/2.

Voting by successive elimination, the Copeland rule, and the Simpson rule all select the Condorcet winner when such a winner exists. Suppose there is no Condorcet winner, and the number of voters is odd. The selection produced by voting by successive elimination then depends on the numbering of the alternatives. The Copeland winner is the alternative that wins the most contests, even if it wins these contests by a small margin and even if it loses some contests by a large margin. The Simpson winner is the alternative whose worst loss is by the smallest margin, even if it does not win many contests overall.

6 Scoring Methods

Voting cycles tend to occur when there is a large disagreement among people. Often in practice there is not too wide a variety of views, and a clear cut Condorcet winner may emerge. Even in such a case, however, one may have reservations about majority rule.

One problem with majority rule is that it may lead to the "tyranny of the majority". A majority of 51% who always agree can impose their will on the minority, who will never get any representation. We will have more to say about this later.

Another problem is that majority rule does not measure the "intensity" of preferences. Consider the following profile:

51%	49%
A	B
B	C
C	D
D	E
E	A

A is the unambiguous winner. Despite the difficulties pointed out above in the coffee/tea/lemonade example, one might still want to make a case for B. No one ranks B lower than second, whereas almost half the people consider A to be the worst possible choice. Put differently, while A barely beats every other alternative, B is unanimously preferred to everything except A, to which it just barely loses. We now consider voting methods that are sensitive to these facts.

A committee must decide among four different locations for a town's dumping site. Under the **Borda count**, each committee member is asked to rank his or her choices from top to bottom. Her top choice is then assigned four points, her next choice three points, her next two points and her least favorite one point. The points given to each alternative by each person are then summed, and the alternatives are ranked according to these totals. The following example illustrates this. There are seven committee members, and they divide into three groups:

3 members	2 members	2 members
C	B	A
B	A	D
A	D	C
D	C	B

A receives a total of 20 points (3x2 + 2x3 + 2x4), B receives a total of 19 points, C gets 18 points, and D gets 13 points. Thus, the Borda count ranks the alternatives:

$$A$$
$$B$$
$$C$$
$$D$$

The Borda Count is one example of a **scoring method**. A scoring method is defined as follows. Suppose there are M alternatives and that each individual has a ranking of them from most to least preferred. To each voter's top choice assign a score of S_1, to the second choice a score of S_2, to the third choice S_3, and so forth up the last choice which receives a score of S_M, where $S_1 \geq S_2 \geq S_3 \cdots \geq S_M$ and $S_1 > S_M$. The total number of points received by each alternative is determined, and the alternatives are ranked by these totals.

The Borda Count is one example of a scoring method, where $S_1 = M$, $S_2 = M - 1$,..., $S_M = 1$. Another example of a scoring method is plurality voting, where $S_1 = 1$, $S_2 = S_3 = \cdots S_M = 0$. Scoring methods have some attractive properties.

Consider a voting body that is divided into two groups, an Upper House and a Lower House. Suppose that in choosing among various bills in separate elections each House chooses the same bill. Then in a joint session we would certainly expect that bill still to be chosen. However, under sequential majority rule this may not be the case. For concreteness suppose there are 24 people in the Upper House, and 28 people in the Lower House. In the Upper House, the representatives fall into three groups with the following rankings

:

8 reps.	8 reps.	8 reps.
Bill A	Bill B	Bill C
Bill B	Bill C	Bill A
Bill C	Bill A	Bill B

The bills are voted on in alphabetical order. First Bill A beats Bill B, then Bill C beats Bill A. C is declared the winner.

The Lower House falls into two groups:

16 reps.	12 reps.
Bill C	Bill B
Bill B	Bill C
Bill A	Bill A

Bill C is ranked first by an absolute majority, so C wins in the Lower House also. Now let us see what happens when we put the two groups together. In the first round Bill Beat B beats Bill A (36 people vote for B, 16 for A). In the second round B wins again (28 vote for B, 24 for C). Although separately both Houses chose Bill C, in a joint session they choose Bill B! [6]

A decision procedure is called **reinforcing** if when two separate groups choose a certain alternative, that alternative is also chosen when the two groups are combined. All scoring methods are reinforcing.

Now consider a vote with 3 bills and 5 voters. The voters divide into four groups with the following preferences:

1 voter	2 voters	2 voters
C	A	B
B	C	A
A	B	C

First Bill B beats A, then C beats B. Now suppose that *two* more voters show up with the following ranking:

$$C$$
$$A$$
$$B$$

[6] Note that when the Lower House is voting alone, we have a voting cycle, which is avoided, as it might be in practice, by having the bills voted on in some specific order.

Now Bill A beats B in the first round, and A beats C in the second round. Look what happens. When the two additional voters show up, A is selected instead of C. But those two voters prefer C to A. They would have been better off staying home!

This is known as the *no-show paradox*. All scoring methods (where any ties are broken according to some fixed ordering) avoid this paradox. On the other hand, if there are at least four alternatives, all Condorcet consistent methods are subject to the no-show paradox and may be non-reinforcing.

Thus, scoring methods have some nice features. As with all unanimous, nondictatorial voting methods, however, they suffer from a dependence on irrelevant alternatives. How serious is this problem? Let us reconsider the committee decision problem we examined at the beginning of this section. Site A was the winning location, with site B receiving the next most points followed by sites C and D. Now suppose that the federal government decrees that site D cannot be used for dumping. Since D was the loser anyway, removing it from consideration will not have much effect on the final outcome, right? Wrong! Look what happens when D is removed. Individuals still rank the remaining sites as before:

3 members	2 members	2 members
C	B	A
B	A	C
A	C	B

But now when we recompute our Borda Count, C emerges as the winner with 15 (3x3 + 2x1 + 2x2) points. In fact, our previous winner, site A, is now ranked last. So how much confidence can we have in our prior conclusion that A was the most preferred site among the alternatives, when a seemingly minor change like dropping the losing site from consideration causes A to be ranked last?

Was this bizarre occurrence just an accident? Well, yes and no. Yes, since it depended upon the particular preference orderings of our seven committee members. For some other rankings, deleting the worst alternative could have no effect whatsoever.

But it was no accident in the following striking manner. Take any number of alternatives and order them in any way (say A, B, C, D, E, F, G). Now delete any alternative at all, and reorder the remaining alternatives any which way (say we delete F, and reorder the remainders E, C, G, D, B, A). Given any scoring method, we can find an example of voters and voter preferences such that the ordering of all the alternatives is the first one, while the ordering when one alternative is deleted is the second one. Put differently, suppose that someone told you the result of an election using a Borda count, and asked you what you could say about how the candidates would be ranked if one of them dropped out. The answer is you could say nothing at all. Anything is possible!

This is true not just for the Borda count, but for any scoring method whatsoever. One of the most common scoring methods is plurality voting. Recall the South Korean presidential elections discussed in the prologue, where the presumed preference profile was:

36.5%	27%	26%
Roh	Kim D.J.	Kim J.P
?	Kim J.P	Kim D.J.
?	Roh	Roh

Just as in the Borda Count example, removing the losing candidate reverses the ranking of the remaining candidates. In fact, the situation is even worse, since either of the losing Kims could beat the winner Roh in a one-on-one election with 53% of the vote. The winner of this plurality election is a **Condorcet loser** who loses to every other candidate in head-to-head contests. An election with five candidates could result in a president who is dispreferred to all four of the other candidates by a majority of the people.

To avoid this last possibility plurality elections are often accompanied by "runoffs". After the first round, the top two vote getters are paired off in a runoff. The (majority) winner of this smaller election is declared the winner. But why take just the top two vote-getters? Why not take the top three vote-getters, have them runoff and then, if still no one has a majority, have the top two vote-getters from this smaller election runoff? Why not start off with the top four vote-getters? Does any of this really matter? It sure does. By changing the runoff procedure we can completely change the outcomes.

With all these problems, there is still one more. A scoring method assigns weights to alternatives according to their positions in voter's rankings. But how should these weights be assigned? If we chose different weights would we get the same rankings (maybe if the weights were not "too" different?) Not at all. By choosing different weights we can get completely different rankings. In fact, there need be no connection whatsoever between the rankings of two different scoring methods, no matter how close these scoring methods appear to be.[7]

7 Restricted Preferences

Arrow addresses the problem of a society attempting to arrive at a social outcome from a collection of distinct individual rankings. Thus, he considers the individual to be the fundamental unit of society. An alternative view would be that people are inextricably linked to the society in which they live, so that we should take as a starting point society viewed as a whole or at least consider certain groups as being fundamental units. Perhaps people do not (or should not) consider themselves as individuals with "individual" preferences but rather take a "social" point of view.

Even if this is the case these social sentiments are in the final analysis held by individuals. Since we have placed no restrictions on individual preferences, these preferences may reflect social feelings, so that, in this respect, there is no real conflict between the individualistic approach and a group oriented approach.

Perhaps, however, the lack of restrictions on individual preferences is not our salvation, but rather the cause of our difficulties. Recall that we require that our decision rule make a selection for any possible realization of individual preferences. Presumably this is because we cannot say anticipate peoples' preferences.

What is the view of society implicit here? We have a group of autonomous individuals who just happen to be in the same place at the same time. Their tastes somehow appear and we have no reason to expect any connection between the desires of any two people.

Is this reasonable? At the very least it seems that people's preferences are shaped by the society in which they live and so are interconnected. A Marxist might further argue that individuals within a certain class will realize their common interests, so that we could for instance, expect all, or most, workers to have the same preferences with respect to many issues.[8]

---

[7] These scoring methods must be truly different, however. For instance, the methods $S_1 = 3$, $S_2 = 2$, $S_3 = 1$; and $S_1 = 6, S_2 = 4, S_3 = 2$, are actually the same method. In the second method, all the numbers have simply been multiplied by two, and this is no real difference. Similarly, adding one to all the numbers would make no real difference, too. In contrast, the scoring method $S_1 = 4$, $S_2 = 2$ $S_3 = 1$, does differ from the above two.

[8] It is no coincidence that Kenneth Arrow comes from the neo-classical economics tradi-

Undoubtedly (or, at the least, plausibly) not all conceivable preference profiles are realizable as a practical matter. Maybe some of the problems we are facing, such as cycling with majority rule, only arise from preference patterns that are themselves not likely to arise.

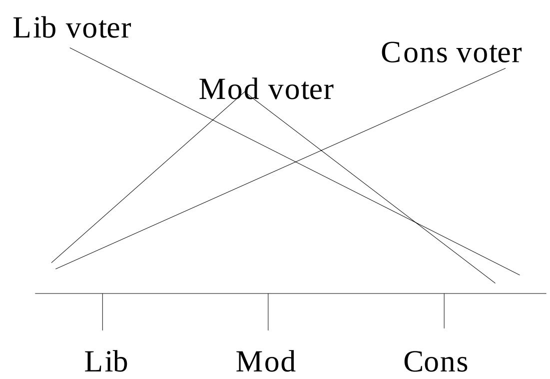
But how are we to decide which patterns are likely to occur? Even if there is a connection between the preferences of different people, what is a reasonable restriction on preference profiles? This is no easy question. The most successful restriction that has been considered is that of single-peaked preferences. This concept is best explained by means of an example. Consider an election between three contestants who can be described as a liberal, a conservative, and a moderate. That is, we can picture these candidates as lying along a line from left to right. Diagrammatically, we have:

$$--Lib.----Mod.---Con.--$$

Electors also divide themselves along the political spectrum. Liberal voters most prefer the liberal candidate, then the moderate candidate, then the conservative candidate. Conservative voters favor the conservative candidate, then the moderate, then the liberal. Moderate voters, most prefer the moderate candidate, then, depending on the voter, either the conservative or the liberal. If we use vertical height to indicate how much an elector likes a candidate, we have a diagram something like the following:

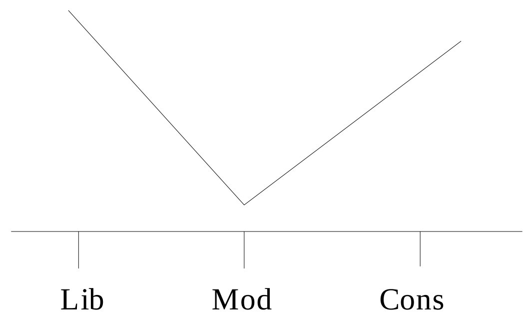
---

tion. The individualistic approach, with exogenously given preferences is a cornerstone of neo-classical economics. Arrow himself, however, is sensitive to critiques of this approach.



Single-Peaked Preferences

This diagram is for three voters and three candidates. Note, that each person's preference is described by a single-peaked graph. Compare this with the preference graph below of a person who likes the conservative most, followed by the liberal, then the moderate:



A Non-Single Peaked Prefernce

There are two (local) peaks in the above diagram, one on the right side and one on the left.

It turns out that if everyone has single-peaked preferences, then pairwise majority rule will always produce an unambiguous winner. Thus, this restriction on preferences circumvents the impossibility theorem.

Theorem 4 *(The Median Voter Theorem): Suppose there are an odd number of voters. Then the preferred alternative of the median voter is a Condorcet winner.*

Though the assumption of single-peakedness has a certain appeal, it is nonetheless quite demanding? Firstly, it requires that all voters (implicitly) agree on a single dimension upon which they can order the alternatives. Secondly, the voters must agree on the alignment of the alternatives along this spectrum. Lastly, this dimension must be of sufficient importance to the voters that they rank the alternatives according to their position on it.

8 Strategic Misrepresentation

Our analysis so far has been concerned with deriving a choice for society from individual rankings. We have seen that such an endeavor is fraught with potential hazards. But there remains one difficulty that we have not even considered yet. How are we to know the individual rankings? We can simply ask people for their rankings, but will they tell the truth? Consider the following successive elimination majority rule example:

Voter I	Voter II	Voter III
A	B	C
B	A	B
C	C	A

First B is put up against C, and B wins; then B beats A as well. At least that is what happens if everybody simply votes according to their preferences. Suppose that voter I knows the preferences of voters II and voters III. Voter I can use this knowledge to manipulate the process to his advantage. Suppose that in the first round I votes for C, although he actually prefers A. For the moment, assume that voters II and III simply continue to vote "sincerely". Then in the first round, C beats B. In the second round, I joins the other players in voting sincerely and A beats C. I has benefitted from this strategic misrepresentation.

Look again at the committee example that began the section on the Borda count. Site A emerged as the winner. But if the two committee members with rankings:

$$\begin{array}{c} B \\ A \\ D \\ C \end{array}$$

were to list their preferences as:

$$\begin{array}{c} B \\ D \\ C \\ A \end{array}$$

site B, their favorite, would win. Again misrepresentation pays off.

A decision rule is strategy-proof if it is always in each individual's interest to truthfully report his or her preferences. As we have just seen, neither voting by successive elimination nor the Borda count is strategy-proof. In fact, nothing reasonable is.

Theorem 5 *(Gibbard-Satterthwaite): Suppose there are at least three alternatives from which to choose and at least two individuals in society. Then the only unanimous, strategy-proof, single-valued decision rule is a dictatorship.*

The 1992 United States Presidential campaign was unusual in that there was an independent party candidate, Ross Perot, with significant support. As the election date approached, however, polls indicated that Perot had virtually no chance of winning. This fact led many people to urge Perot supporters not to "waste" their vote on Perot, but rather to vote for one of the major party candidates. In our terminology, Perot supporters were being urged to vote strategically.

Any rule that we use for group decision making will give somebody sometime the incentive to misrepresent his true preferences. This might concern us for several reasons.

First of all, we might be morally distressed by the fact that our social procedure encourages people to act in a less than forthright manner. Second of all, we might consider it somehow unfair that certain people are able to manipulate the system to their advantage. Third of all, if people are voting strategically, it is difficult to interpret the results of elections. What are we to make of the fact that a winning candidate received 40% of the vote if we do not know how many voters actually preferred a candidate for whom they did not vote?

Finally, the possibility of strategic manipulation creates a complex strategic problem for the players which may result in perverse outcomes. Recall the example with which we began. Player I could manipulate the voting scheme by voting for C in the first round, although he preferred B. This was assuming that voters II and III naively vote according to their true preferences. But why make this assumption? What if instead II and III are as clever as I? Now in the first round I must consider not simply II and III's preferences, but how they are going to vote. How should I decide how II and III are going to vote? Surely this depends on how they think he is going to vote (and how they think each other will vote). And how II, for instance, thinks I is going to vote depends on how II thinks I thinks II is going to vote which depends on how II thinks I thinks II thinks I thinks is going to vote etc... How can such a complex problem be resolved? Is there a definite right way the players should vote? Although game theory has made considerable progress on this problem, it cannot be resolved definitively. In any case, as our next example shows strategizing on the part of voters may have unintended consequences.

Consider a law school hiring a new dean. Three brave souls have put their names up for consideration. The first, CL, is a critical legal theorist, the second, DD, a doctrinalist, and the third, UA, an unemployed artist who figures that being dean of the law school will help pay his rent while he awaits his big break. The faculty divides into two groups; approximately half favour CL, while the rest favour DD. In the (rare) spirit of scholarship, all agree that both CL and DD are better than UA. Thus, we have:

APPROX. 50	APPROX 50
CL	DD
DD	CL
UA	UA

Every faculty member is asked to rank the candidates. A scoring method is used in which each top listing is worth 5 points, each second listing is worth 3 points, and each bottom listing is worth 0 points. Thus, if each professor truthfully lists his or her preferences, either CL or DD will win with a little over 400 points.

But now the supporters of CL get together and decide to be a little clever. They realize that by listing their preferences as:

$$\begin{array}{c} CL \\ UA \\ DD \end{array}$$

Unfortunately, however, DD's supporters are being equally clever. They decide to help their candidate by misreporting their preferences as:

$$\begin{array}{c} DD \\ UA \\ CL \end{array}$$

As a result UA wins the deanship with about 300 points to CL and DD's 250 points. This despite the fact that everybody agrees that both CL and DD are preferable to UA. Thus, although a voting procedure might guarantee that a unanimous loser cannot win when everybody votes sincerely, when people act strategically the procedure may yield a unanimous loser.

The examples of strategic misrepresentation we have just considered suggest people manipulating the decision process to their advantage by subverting society's "true" desires. Thus, in the first example, I manages to get his favorite alternative A selected, although a majority of the people prefer B to A.

Manipulation need not have this subversive aspect, however. Reconsider the 1987 South Korean presidential election. If the supporters of Kim Jong Pil had misrepresented and voted for Kim Dae Jung, this latter would have been elected and would arguably have been a more representative choice than the actual winner.

In the previous section we saw that if we restrict ourselves to single-peaked preferences, Arrow's theorem can be avoided. This restriction also helps here.

Theorem 6 *Suppose that preferences are single-peaked. Then any Condorcet consistent decision rule is strategy-proof.*

9 Agenda Manipulation

10 Conclusion

Arrow's justly famous impossibility theorem is the cornerstone of social choice theory. Essentially, two types of resolutions to this negative result have been proposed. The first type notes that Arrow required that his decision rule be defined for all possible preference profiles. But as a practical matter, not all possible profiles are likely to arise. Perhaps many difficulties are merely "theoretical possibilities" that are not likely to arise, and therefore of not much actual concern.

Work along these lines has not been very encouraging. While single-peakedness does rule out many paradoxes, and may be a good assumption in some situations, there are not many reasonable restrictions that guarantee no problems. It seems fair to say that we should expect any system to exhibit some problems some times.

The second type of resolution focuses on the pair-wise comparison, or independence property. One reaction to Arrow's theorem is that it is neither surprising, nor undesirable, that no decision rule is independent of irrelevant alternatives. Let us examine this response more closely.

An interesting interpretation of the impossibility theorem involves the ranking of decathlon athletes. Think of the athletes as candidates. Think of the events as the voters. Then each event ranks the athletes by how well they finish in that event. Thus, the 100 meter dash most "prefers" the athlete with the smallest running time, then "likes" the second place finisher, etc... The high jump most prefers the highest jumper, and so forth. The impossibility theorem tells us that the winner of the decathlon cannot be determined just by taking the athletes two at a time and seeing who beat whom in the most events.

But why would we want to rank the decathletes in this manner? If one athlete beats another in a photo finish in the 100 meter race, should this balance his loss by six inches in the high jump? Of course not, and this is exactly why the decathlon is not decided in this way (or maybe the founders of the Olympics anticipated Arrow's manuscript).

Or consider a family deciding on its vacation plans. What family would simply ask for pair-wise comparisons of the alternatives without trying to find out how strongly each family member felt? Since in pair-wise comparisons Arrow asks only that each family members give their preferred alternative, he leaves no room for an expression of the intensity of their feelings.

So perhaps we should never have been looking for independence of irrelevant alternatives at all. The problem, as we know amply well by now, is that when we relax Arrow's condition we are asking for trouble. For instance, how are we to compare a finish in the 100 meter dash with a finish in the high jump? With equally reasonable procedures we can completely reverse the rankings of athletes.

So the impossibility theorem is a true dilemma, although perhaps it should come as no surprise. After all, given the range of views in society, how could we possibly hope to amalgamate these views in a trouble-free way? In a sense, the statement that we cannot easily combine preferences is only to be expected.

This may be, but the power of Arrow's work is two-fold. First of all, he manages to give a precise meaning to the rather vague feeling that there will be "difficulties" in aggregating individual preferences. Second of all, he shows just how fundamental these difficulties are.

11 Some References

Most of the material here is taken from other sources. Arrow's impossibility theorem can be found in his monograph Social Choice and Individual Values. Perhaps a more useful secondary source is Sen's book Collective Choice and

Social Welfare. The book is conveniently divided into technical and non-technical sections. Both these books follow Arrow's original approach of looking for a social welfare function which provides a complete ranking of all other alternatives, rather than a decision rule which seeks only a top choice.

Two excellent but quite difficult books are Moulin's The Strategy of Social Choice and Axioms of Cooperative Decision Making. Somewhat dated, but possibly still the best source, is Luce and Raiffa's classic text Games and Decisions.

Theorem 2 is due to Muller and Satterthwaite, Theorem is due5 to Gibbard and Satterthwaite, and Theorem 3 is due to May. The results on scoring methods are from Young, Saari and Fishburn. The no-show paradox is due to Brams.

12 Proofs

In this section we provide rigorous proofs of theorems 1, 2, and 5. The mathematics involved are not hard, although some of the arguments are fairly subtle. For the proofs we assume that no individual is indifferent between any two alternatives, although the theorems are true even if we allow indifference.

Proof of Theorem 1: (Reductio ad absurdum) Suppose we have a unanimous, non-dictatorial, pair-wise decision rule. The proof proceeds in two steps.

1) Call a set of individuals decisive for A against B, denoted D(A,B), if A is selected in a pair-wise contest with B whenever they all rank A above B.

Call a set of individuals nearly decisive for A against B, denoted ND(A,B), if A is selected above B whenever they all rank A above B, and everyone else ranks B above A.

We first prove that if some individual, I, is nearly decisive for some pair of alternatives, say for A against B, then that individual is a dictator. Accordingly, suppose that I is nearly decisive for A against B, and that preferences are as follows:

I	OTHERS
A	B
B	A or C
C	C or A

Notice that for people other than I, we have not specified how A compares to C. Since I is nearly decisive for A against B, A must beat B in a pair-wise contest. Since B is unanimously preferred to C, B must beat C pair-wise. Therefore, only A can be selected as on overall winner that is also a pair-wise winner. In particular, we have that A beats C. We have just shown that in a pair-wise contest between A and C, if I prefers A to C, then regardless of how others feel about A and C, A will be chosen (of course, we have so far assumed that I prefers A to B and that others prefer B to A, but in a contest between A and C this information is irrelevant (indeed, "non-existent") so this cannot have had an effect upon our conclusion). Thus, I is decisive for A against C.

Now suppose that we have the following preferences:

I	OTHERS		
C	C		B
A	B	or	C
B	A		A

Since C is unanimously preferred to A, C must be chosen in a contest against A. Since I is nearly decisive for A against B, A beats B. Therefore, C must be chosen above A and B, and in particular C beats B. Again, we have only stipulated that I prefers C to B, so I is decisive for C against B.

So far we have shown that ND(A,B) => D(A,C) and D(C,B). But since D => ND, by simply changing labels we also have:

D(A,C) => D(A,B) and D(B,C)

D(B,C) => D(B,A)

D(C,B) => D(C,A)

Thus, we have that I's preferences completely determine which outcome is selected from any pair that includes A or B. But from the (derived) conclusion ND(A,C) we similarly conclude that I determines which outcome is selected from any pair that includes (the arbitrarily) chosen C. In summary, from the assumption that I is nearly decisive for some pair A and B, we have concluded that for any pair of alternatives whatsoever, I's preferences determine which outcome is selected. Since the decision rule must be consistent with pair-wise comparisons, I's preferences determine the outcome and I is a dictator.

2) Since the decision rule is unanimous, the set of all individuals is nearly decisive for any two alternatives. Now consider the smallest set V that is nearly decisive for some alternatives. Suppose these alternatives are A against B. If V is a single individual, then from 1) above, that individual is

a dictator. Therefore, suppose that V has at least two individuals, and split V into two sets, V_1 with just one individual, and V_2 with the rest. Let V_3 (possibly empty) be all those people not in V.

Now consider the following preferences:

V_1	V_2	V_3
A	C	B
B	A	C
C	B	A

A must be chosen above B since V is nearly decisive for A against B and everyone in V prefers A to B, while all others prefer B to A. If C is chosen over B, then V_2 is nearly decisive. Since V_2 is smaller than V, the smallest decisive set, this is impossible. Hence, C is at best tied with B, and the overall winner must be A. But this makes V_1 nearly decisive for A against C, which is also impossible.

Therefore, our original assumption that we had an appropriate decision rule must be false.

Q.E.D.

Theorem 2

We say that the position of an alternative **improves** in going from one set of rankings to another if the only difference between the two sets is that in the latter some people rank that alternative higher than they did previously.

We say that a decision rule is **strongly monotonic** if an improvement in the position of some alternative either causes that alternative to be selected, or has no effect on the outcome. Clearly, if a decision rule is strongly monotonic, a disimprovement in the position of an alternative which is not selected can have no effect on the outcome.

Theorem **2** says that the only efficient, strongly monotonic decision rule is a dictatorship.

Proof of Theorem 2

Suppose we have an efficient, strongly monotonic decision rule defined over N alternatives that selects some alternative A given some profile. We use this rule to define a pair-wise decision rule as follows. Given A and some other alternative, say B, fix the individual rankings of A and B, and now for each person, lower the positions of all the other alternatives until they are below A and B, in any order whatsoever. Note:

1) By strong monotonicity, the decision rule must still select A.

2) A is being selected over B, given the initial rankings of A and B, but completely irrespective of the original rankings of the other alternatives. Thus, this defines a pair-wise decision rule (given the initial preference profile, for a paired comparison of any alternatives x and y, select the alternative that the rule selects when all the other alternatives are pushed below x and y in everybody's rankings (by unanimity and stong monotonicity, either x or y will be chosen)).

3) From 1) and 2) we have a unanimous decision rule such that the selected outcome also beats all other alternatives in pair-wise comparisons. From Theorem 1, this rule must be a dictatorship.

Q.E.D.

Proof of Theorem 5:

Suppose that we have an efficient, strategy proof decision rule, and that for some profile of preferences outcome A is selected. Now improve the position of some alternative C, which may or not be the same as A, in someone's ranking. Suppose that outcome B is now chosen. Note that if the person for whom the position of C has improved had not reported this change, A would still have been chosen. Therefore, for this person to be willing to truthfully report this change, B must be at least as good as A. On the other hand, for this person not to have wanted to falsely claim this improvement beforehand, B must have been no better than A before. One possibility is that B is A. Otherwise, since only the position of C has changed relative to A, it is the only alternative that can be better than A now, although it was worse than A before. Therefore, B must be either C or A. This means that the decision rule is strongly monotonic and hence, from Theorem 2, a dictatorship.

Q.E.D.