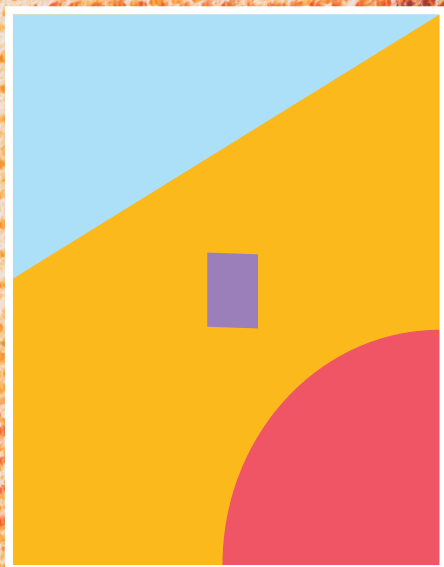


Models in Microeconomic Theory

Expanded Second Edition

Martin J. Osborne
Ariel Rubinstein





<https://www.openbookpublishers.com>

© 2020, 2023 (expanded second edition) Martin J. Osborne and Ariel Rubinstein



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs license (CC BY-NC-ND 4.0). This license allows you to share, copy, distribute, and transmit the work providing you do not modify the work, you do not use the work for commercial purposes, you attribute the work to the authors, and you provide a link to the license. Attribution should not in any way suggest that the authors endorse you or your use of the work and should include the following information.

Martin J. Osborne and Ariel Rubinstein, *Models in Microeconomic Theory* (Expanded Second Edition), Cambridge, UK: Open Book Publishers, 2023, <https://doi.org/10.11647/OBP.0361>.

To access detailed and updated information on the license, please visit <https://www.openbookpublishers.com/product/1171#copyright>.

For further details on CC BY-NC-ND licenses, see <https://creativecommons.org/licenses/by-nc-nd/4.0/>. All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at <https://archive.org/web>.

Updated digital material and resources associated with this book are available at <https://www.openbookpublishers.com/product/1171#resources>.

Every effort has been made to identify and contact copyright holders and any omission or error will be corrected if notification is made to the publisher.

ISBN Paperback: 978-1-80511-121-4

ISBN Hardback: 978-1-80511-122-1

ISBN Digital (PDF): 978-1-80511-123-8

DOI: 10.11647/OBP0361

This book is available in two versions, one that uses feminine pronouns and one that uses masculine pronouns. This version uses feminine pronouns.

Version: 2023.5.30 (s)

20 Aggregating preferences

When we discuss public decisions, we often talk about the preferences of a group of people, like a nation, a class, or a family. We do so even though the members of the group have different preferences; we say that “the group prefers one option to another” even though the meaning of such a statement is unclear. In this chapter we discuss one model of the aggregation of individuals’ preferences into a social preference relation.

20.1 Social preferences

Society consists of a set of individuals, each of whom has preferences over a set of social alternatives X , which is fixed throughout our discussion. The information we have about each individual’s preferences is purely ordinal, in the sense that it tells us only that the individual prefers one alternative to another, or regards two alternatives as indifferent. In particular, it does not specify the intensity of an individual’s preferences. On the basis of the information, we cannot say, for example, that an individual “prefers a to b more than she prefers c to d ” or that “individual i prefers a to b more than individual j prefers b to a ”.

We want to aggregate the individuals’ preferences into a social preference. A voting procedure is sometimes used to do so. An example is the method used in the Eurovision song contest. In this case X is the set of songs performed in the contest. Each country submits an ordered list of the members of X that expresses its preference ordering (which itself is an aggregation of the orderings of the listeners in that country). The countries’ rankings are aggregated by assigning points to the songs according to their positions in each country’s ranking (12 for the top song, 10 for the second, 8, 7, ..., 1 for the next 8 songs, and 0 for all others) and then summing the points across countries to give the European ranking. This method is a special case of a family of aggregation methods called scoring rules, which we define later.

Majority rule is a simpler, and natural, principle for determining a social preference. According to this rule, the society prefers alternative a to alternative b if a majority of individuals prefer a to b . A major difficulty with this rule is that when X contains more than two alternatives, the resulting binary relation may not be **transitive**, so that it is not a **preference relation**. Recall the **Condorcet paradox**

discussed in [Chapter 1](#). A society contains three individuals, 1, 2, and 3, and three social alternatives, a , b , and c . The individuals' preferences are given in the following table, where each column lists the alternatives in the order of one individual's preferences.

1	2	3
a	b	c
b	c	a
c	a	b

Individuals 1 and 3, a majority, prefer a to b , individuals 1 and 2 prefer b to c , and individuals 2 and 3 prefer c to a . Thus the binary relation \succsim defined by majority rule satisfies $a \succsim b \succsim c \succsim a$, and hence is not transitive.

20.2 Preference aggregation functions

The central concept in this chapter is a preference aggregation function (PAF), which maps the preferences of the individuals in a society into a single “social” preference relation. A PAF is usually called a social welfare function. We avoid this term because the concept is not related to the individuals' welfare in the everyday sense of that word.

Definition 20.1: Society

A *society* consists of a finite set N (the set of *individuals*) and a finite set X (the set of *alternatives*).

We assume that every individual in N has a [preference relation](#) over X . For simplicity we assume that each of these preference relations is [strict](#) (no individual is indifferent between any two alternatives). A profile of preferences specifies a (strict) preference relation for every individual.

The domain of a PAF is the set of all preference profiles. That is, we do not impose any restrictions on the set of preference profiles. Problem 2 illustrates the significance of this assumption. To each preference profile, a PAF assigns a single preference relation. We do not assume that this preference relation is strict: the social preference may have indifferences.

Definition 20.2: Preference aggregation function

A *preference aggregation function* (PAF) for a [society](#) $\langle N, X \rangle$ is a function that assigns a (“social”) [preference relation](#) over X to every profile of strict [preference relations](#) over X .

Note that a preference aggregation function is not a single preference relation. Rather, it is a rule for aggregating the individuals' preference relations into a single preference relation. Here are a few examples.

Example 20.1: Counting favorites

The alternative x is ranked above y if the number of individuals for whom x is the best alternative is greater than the number of individuals for whom y is the best alternative. Formally, for any preference profile $(\succsim^i)_{i \in N}$, the social preference relation \succsim is defined by

$$x \succsim y \text{ if } |\{i \in N : x \succ^i a \text{ for all } a \in X\}| \geq |\{i \in N : y \succ^i a \text{ for all } a \in X\}|.$$

(Note that every alternative that is not any individual's favorite is indifferent to any other such alternative in the social preferences, and is ranked below all other alternatives.)

Example 20.2: Scoring rules

Given a preference relation \succsim^i , the *position* $K(\succsim^i, x)$ of alternative x in \succsim^i is the number of alternatives that are at least as good as x according to \succsim^i : $K(\succsim^i, x) = |\{a \in X : a \succsim^i x\}|$. Thus the position of the best alternative according to \succsim^i is 1, the position of the second-ranked alternative is 2, and so on. A scoring rule is characterized by a function p that gives the number of points $p(k)$ credited to an alternative for being ranked in each position k . Assume $p(1) \geq \dots \geq p(|X|)$. Alternatives are compared according to the sum of the points they accumulate across the individuals' preferences. Precisely, the scoring rule defined by p maps the preference profile $(\succsim^i)_{i \in N}$ into the social preference relation \succsim defined by

$$x \succsim y \text{ if } \sum_{i \in N} p(K(\succsim^i, x)) \geq \sum_{i \in N} p(K(\succsim^i, y)).$$

Note that the previous example (Example 20.1) is the scoring rule for which p is given by $p(1) = 1$ and $p(k) = 0$ for all $k \geq 2$.

Example 20.3: Pairwise contests

For every pair (x, y) of distinct alternatives, say that x beats y if a majority of individuals prefer x to y , and x and y tie if the number of individuals who prefer x to y is the same as the number of individuals who prefer y to x . Assign to each alternative one point for every alternative it beats and half a point for every alternative with which it ties. Rank the alternatives by the total number of points received.

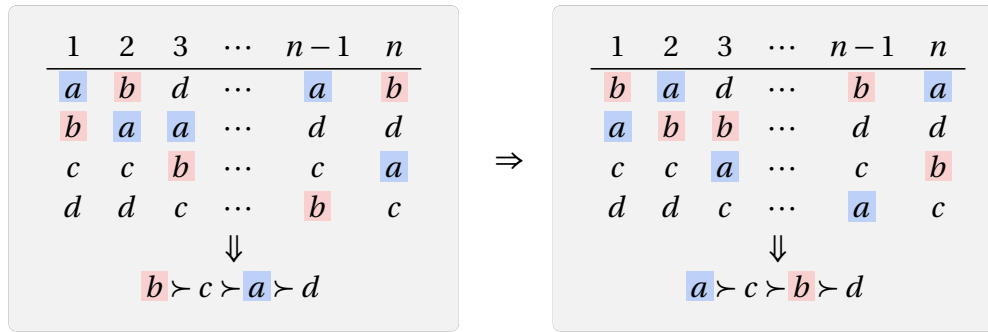


Figure 20.1 An illustration of the **neutrality property**. Each column shows the preference ordering of the individual whose name heads the column, from best at the top to worst at the bottom. When the alternatives a and b are interchanged in every individual's preferences, they are interchanged also in the social preferences.

The last two examples are extreme.

Example 20.4: External preferences

The social preference relation is some fixed given preference relation regardless of the individuals' preferences. That is, for some preference relation \succ^* we have $x \succ y$ if and only if $x \succ^* y$.

Example 20.5: Dictatorship

The social preference relation coincides with the preference relation of one of the individuals, called the dictator. That is, for some individual i^* we have $x \succ y$ if and only if $x \succ^{i^*} y$.

20.3 Properties of preference aggregation functions

How do we evaluate a preference aggregation function? We proceed by specifying properties that seem appealing, and then look for preference aggregation functions that satisfy them.

Neutrality A PAF is neutral if it treats the alternatives symmetrically. That is, for any preference profile, if we interchange the positions of any pair of alternatives in every individual's preference relation then the PAF responds to the change by interchanging the positions of these alternatives in the social preference relation. The property is illustrated in [Figure 20.1](#).

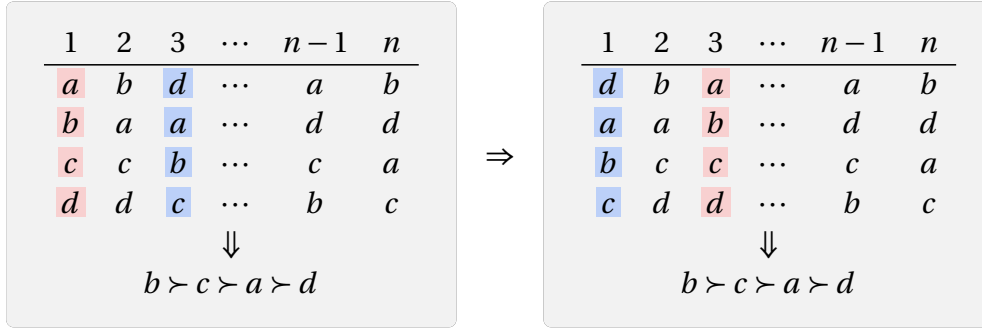


Figure 20.2 An illustration of the [anonymity property](#). When the preferences of two individuals are interchanged, the social preferences remain the same.

Definition 20.3: Neutrality

A [preference aggregation function](#) F for a [society](#) $\langle N, X \rangle$ is *neutral* if for any profiles $(\succsim^i)_{i \in N}$ and $(\succeq^i)_{i \in N}$ of preference relations over X such that for some alternatives x and y the preference relation \succeq^i for each $i \in N$ is obtained from \succsim^i by interchanging the positions of x and y , then the social preference relation $F((\succeq^i)_{i \in N})$ is obtained from $F((\succsim^i)_{i \in N})$ by interchanging the positions of x and y as well.

For $X = \{a, b\}$, an example of a PAF that is not neutral is the one that ranks a above b only if at least $\frac{2}{3}$ of the individuals prefer a to b . All the examples in the previous section except [Example 20.4 \(external preferences\)](#) are neutral.

Anonymity A PAF is anonymous if it does not discriminate between individuals. Consider two profiles in which all preference relations are the same except those of i and j , which are interchanged. An anonymous PAF generates the same social preference relation for these two profiles. The property is illustrated in [Figure 20.2](#). All the examples in the previous section except [dictatorship](#) are anonymous.

Definition 20.4: Anonymity

A [preference aggregation function](#) F for a [society](#) $\langle N, X \rangle$ is *anonymous* if whenever $(\succsim^i)_{i \in N}$ and $(\succeq^i)_{i \in N}$ are profiles of preference relations over X with $\succsim^j = \succeq^k$, $\succsim^k = \succeq^j$, and $\succsim^i = \succeq^i$ for all $i \in N \setminus \{j, k\}$, we have $F((\succsim^i)_{i \in N}) = F((\succeq^i)_{i \in N})$.

Positive responsiveness A PAF is positively responsive if whenever an alternative x rises one step in one individual's preferences and all other individuals'

preferences remain the same, any alternative z that was originally ranked no higher than x in the social preferences is now ranked below x in these preferences. The examples in the previous section are positively responsive except for [Example 20.4](#) when the external preference relation has indifferences.

Definition 20.5: Positive responsiveness

A **preference aggregation function** F for a **society** $\langle N, X \rangle$ is *positively responsive* if for any profile $(\succsim^i)_{i \in N}$ of preference relations over X and any profile $(\succeq^i)_{i \in N}$ that differs from $(\succsim^i)_{i \in N}$ only in that there exist $j \in N$, $x \in X$, and $y \in X$ such that $y \succ^j x$ and $x \succeq^j y$ (i.e. x rises one step in j 's ranking), then for any $z \in X$ with $x \succsim z$ we have $x \succ z$, where $\succsim = F((\succsim^i)_{i \in N})$ and $\succeq = F((\succeq^i)_{i \in N})$ (and \succ is the strict relation derived from \succeq).

Pareto A PAF satisfies the Pareto property if for any preference profile in which all individuals rank x above y the social ranking ranks x above y .

Definition 20.6: Pareto property

A **preference aggregation function** F for a **society** $\langle N, X \rangle$ satisfies the *Pareto property* if whenever $x \succ^i y$ for all $i \in N$ for a profile $(\succsim^i)_{i \in N}$ of preference relations over X we have $x \succ y$, where $\succsim = F((\succsim^i)_{i \in N})$.

A scoring rule ([Example 20.2](#)) satisfies this property if the weighting function p is decreasing (rather than merely nonincreasing). [Example 20.4](#) (external preferences) does not satisfy the property.

Independence of irrelevant alternatives The last property we define is independence of irrelevant alternatives (IIA). This property says that the social ranking of any alternatives a relative to b depends only on the individuals' rankings of a relative to b and not on their rankings of any other alternatives or on their rankings of a or b relative to any other alternative.

Definition 20.7: Independence of irrelevant alternatives

A **preference aggregation function** F for a **society** $\langle N, X \rangle$ is *independent of irrelevant alternatives* (IIA) if for any profiles $(\succsim^i)_{i \in N}$ and $(\succeq^i)_{i \in N}$ of preference relations over X for which there are alternatives x and y with

$$x \succ^i y \text{ if and only if } x \succeq^i y \text{ for every } i$$

we have

$$x \succ y \text{ if and only if } x \succeq y$$

where $\succsim = F((\succsim^i)_{i \in N})$ and $\succeq = F((\succeq^i)_{i \in N})$.

Dictatorship (Example 20.5) satisfies this property: the dictator's ranking of any two alternatives determines the social ranking of these alternatives. Some scoring rules (Example 20.2) do not satisfy the property. For example, consider a society with three individuals and three alternatives. Assume that the weighting function p for the scoring rule satisfies $p(1) - p(2) > 2(p(2) - p(3)) > 0$ and consider the following two preference profiles.

1	2	3
c	c	c
b	a	a
a	b	b

1	2	3
b	c	c
c	a	a
a	b	b

The relative rankings of a and b in the two profiles are the same, but the scoring rule ranks a above b for the left profile and b above a for the right profile.

20.4 Arrow's impossibility theorem

The central result of this chapter says that the only PAFs that satisfy the **Pareto property** and **IIA** are **dictatorships**. That is, the requirements that the social comparison of any two alternatives depends only on the individuals' binary comparisons of these two alternatives, and not on their preferences over other pairs, and that one alternative is socially preferred to another whenever there is consensus, do not leave room for real aggregation of the individuals' preferences.

Proposition 20.1: Arrow's impossibility theorem

Let $\langle N, X \rangle$ be a **society** for which X contains at least three alternatives. A **preference aggregation function** F satisfies the **Pareto property** and **IIA** if and only if it is **dictatorial**: there is an individual $i^* \in N$ such that for every profile $(\succsim^i)_{i \in N}$ of preference relations over X we have $x \succsim y$ if and only if $x \succsim^{i^*} y$, where \succsim is the social preference relation $F((\succsim^i)_{i \in N})$.

Proof

If F is dictatorial then it satisfies the Pareto property and IIA. We now show the converse.

Let F be a PAF that satisfies the Pareto property and IIA. Fix $b \in X$.

Step 1 Consider a preference profile for which b is either at the top or the bottom of each individual's ranking and let \succsim be the social preferences attached to the profile by F . For such a profile, b is either the unique \succsim -maximal alternative or the unique \succsim -minimal alternative.

Proof. Assume to the contrary that for two other alternatives a and c we have $a \succsim b \succsim c$. Consider a preference profile that is obtained from the original profile by moving c just above a for every individual who originally prefers a to c (so that it remains below b for all individuals for whom b is best), as illustrated below.

1	2	3	...	$n-1$	n		1	2	3	...	$n-1$	n
b	b		...	a	b		b	b		...	c	b
	a	a	...			\rightarrow		c	c	...	a	
c	c		...	c	a		c	a	a	...		c
		c	...		c					...		a
a		b	...	b			a		b	...	b	

Denote by \succeq the social preference relation that F generates for the new preference profile. By the **Pareto property**, $c \succ a$. By **IIA**, the rankings of a and b remain unchanged, so $a \succeq b$, and the rankings of b and c remain unchanged, so $b \succeq c$. Thus $a \succeq c$ by transitivity, contradicting $c \succ a$. \triangleleft

Step 2 Consider two preference profiles in which b is either at the top or the bottom of each individual's ranking and in which the set of individuals who rank b at the top is the same. Let \succsim and \succeq be the social preferences attached by F to the profiles. Then either b is the unique maximal alternative for both \succsim and \succeq or is the unique minimal alternative for both \succsim and \succeq .

Proof. Consider one of the profiles. By Step 1, \succsim either ranks b uniquely at the top or uniquely at the bottom. Suppose that it ranks it at the top. The ranking of b relative to any other alternative x is the same in both profiles and hence by **IIA** b and x are ranked in the same way by both \succsim and \succeq . Thus b is ranked at the top of \succeq . If b is ranked at the bottom of \succsim the argument is analogous. \triangleleft

Step 3 For some individual i^* ,

- i. for every preference profile for which $1, \dots, i^* - 1$ rank b at the top and i^*, \dots, n rank it at the bottom, the preference relation attached by F to the profile ranks b uniquely at the bottom

- ii. for every preference profile for which $1, \dots, i^*$ rank b at the top and $i^* + 1, \dots, n$ rank it at the bottom, the preference relation attached by F to the profile ranks b uniquely at the top.

Proof. Take a preference profile in which b is at the bottom of all individuals' preferences. By the [Pareto property](#), b is the unique minimal alternative for the attached social preferences. Now, for each individual in turn, starting with individual 1, move b from the bottom to the top of that individual's preferences. By Step 1, b is always either the unique maximal or unique minimal alternative in the social preferences. By the [Pareto property](#), it is the unique maximal alternative of the attached social preferences after it moves to the top of all individuals' preferences. Thus for some individual i^* the change in her preferences moves b from the bottom to the top of the social preferences. By Step 2 the identity of i^* does not depend on the individuals' rankings of the other alternatives. \triangleleft

Step 4 For all alternatives a and c different from b we have $a \succ c$ if and only if $a \succ^{i^*} c$.

Proof. Assume to the contrary that for some preference profile we have $a \succ^{i^*} c$ and $c \succ a$. Modify the profile by raising b to the top of the preferences of individuals $1, \dots, i^* - 1$, lowering it to the bottom of the preferences of individuals $i^* + 1, \dots, n$, and moving it between a and c for individual i^* , as in the following illustration.

1	2	...	i^*	...	$n-1$	n		1	2	...	i^*	...	$n-1$	n
	c			...	a	b		b	b	a	
b		...	a	...			\rightarrow		c		a	...		c
a		c	c		a		...	b	...	c	a
	a	...	c	...	b	a					
c	b		b	...				c	a		c	...	b	b

Denote by \succeq the social preference that F attaches to the new preference profile. The relative positions of a and c are the same in the two profiles, so $c \succeq a$ by [IIA](#). In the new profile, the individuals' rankings of a relative to b are the same as they are in any profile in which b is ranked at the top by $1, \dots, i^* - 1$ and at the bottom by the remaining individuals, so that by Step 3 and [IIA](#) we have $a \succ b$. Similarly, $b \succ c$. Thus by transitivity $a \succ c$, a contradiction. \triangleleft

Step 4 states that i^* is the dictator regarding any two alternatives other than b . It remains to show that i^* is also the dictator regarding the comparison of b with any other alternative.

Step 5 For every alternative a we have $a \succ b$ if and only if $a \succ^{i^*} b$.

Proof. Consider a preference profile for which $a \succ^{i^*} b$. Let c be an arbitrary third alternative. Modify the profile by moving c in i^* 's ranking to between b and a (if it is not already there) and raising c to the top of all the other individuals' rankings, as in the following illustration.

1	...	i^*	...	$n-1$	n		1	...	i^*	...	$n-1$	n
c			...	a	b	→	c			...	c	c
b	...	a	...				b	...	a	...	a	b
a	...	b	...	c	c		a	...	c	...		
	b	a				b	...	b	a
		c		

Denote by \succeq the social preference attached by F to the new profile. By the **Pareto property**, $c \succ b$. By **Step 4**, i^* determines the social preference between a and c , so that $a \succ c$, and hence $a \succ b$ by transitivity. Since the relative preferences of a and b are the same in the two profiles, we have also $a \succ b$ by **IIA**. ◁

Comments

1. So far we have interpreted a PAF as a method of generating a social preference relation from the profile of the individuals' preferences. We can alternatively think of a PAF as a method for an individual to form preferences over a set of objects on the basis of several criteria. For example, an individual may base her preferences over cars on their relative price, their ranking by a car magazine, and their fuel consumption. In this context, **Proposition 20.1** says that if the preference-formation process satisfies the Pareto property and the preference between any two alternatives is a function only of the way the criteria rank them, then the generated preference relation is intransitive unless it coincides with the ranking by one of the criteria.
2. **Proposition 20.1** is interpreted by some as a proof of the impossibility of aggregating the preferences of the members of a group. This interpretation is incorrect. At most the result says that aggregating preferences in such a way that the social ranking of any two alternatives depends only on their

relative rankings by the individuals is not possible. This requirement can be viewed as a simplicity constraint on the aggregation process. [Scoring rules](#) with decreasing weighting functions satisfy all the properties discussed in this chapter except IIA.

3. The requirement that a PAF is defined for all possible preference profiles is very demanding. In many contexts only some preference profiles make sense. In fact, for some meaningful restrictions on the set of preference profiles, majority rule induces a transitive relation. Suppose that the alternatives are ordered along a line (from left to right on the political spectrum, for example) each individual has a favorite alternative, and each individual's ranking falls away on each side of this favorite, so that her preferences are single-peaked. In this case, majority rule induces a social preference relation (see Problem 2).
4. As we mentioned earlier, a preference aggregation function uses information only about the individuals' ordinal rankings; it does not take into account the individuals' intensities of preference and does not compare these intensities across individuals. Consider a society with two individuals and two alternatives. If one individual prefers a to b while the other prefers b to a , then a reasonable assessment of the individuals' aggregated preference would compare the degree to which individual 1 likes a better than b with the degree to which individual 2 likes b better than a . Such information is missing from the model.

20.5 Gibbard-Satterthwaite theorem

We close the chapter with another classical result, the Gibbard-Satterthwaite theorem. This result involves the concept of a social choice rule, which is related to, but different from, a preference aggregation function. Whereas a preference aggregation function assigns a preference relation to each preference profile, a social choice rule assigns an *alternative* to each preference profile, interpreted as the alternative to be chosen given the individuals' preferences.

Definition 20.8: Social choice rule

A *social choice rule* for a [society](#) $\langle N, X \rangle$ is a function that assigns a member of X (an alternative) to every profile of strict [preference relations](#) over X .

We are interested in social choice rules that satisfy two properties. The first one requires that a social choice rule selects an alternative if there is a consensus that it is the best alternative.

Definition 20.9: Unanimous social choice rule

A **social choice rule** f for a **society** $\langle N, X \rangle$ is *unanimous* if for every profile $(\succsim^i)_{i \in N}$ of preference relations over X , $f((\succsim^i)_{i \in N}) = x$ if x is \succsim^i -optimal for all $i \in N$.

The second property is central to the result. Imagine that the social planner asks the individuals about their preferences. The property requires that for each individual, whatever her true preference relation, reporting this preference relation is optimal regardless of the other individuals' reports. This property, called strategy-proofness, is discussed in Chapter 17 (Definition 17.4).

Definition 20.10: Strategy-proof social choice rule

The **social choice rule** f for a **society** $\langle N, X \rangle$ is *strategy-proof* if for every individual $j \in N$ and every profile $(\succsim^i)_{i \in N}$ of preference relations over X , we have $f((\succsim^i)_{i \in N}) \succsim^j f((\succeq^i)_{i \in N})$ for any preference relation \succeq^j of individual j , where $\succeq^i = \succsim^i$ for all $i \in N \setminus \{j\}$. That is, j optimally reports her true preference relation regardless of the other individuals' reports.

Which social choice rules are unanimous and strategy-proof? The striking answer is that if X contains at least three alternatives then the only such rules are dictatorships, for which there is an individual (the dictator) whose top reported alternative is always chosen. Any dictatorship is strategy-proof because the dictator cannot gain by misreporting her preferences, given that her top reported alternative is chosen, and no other individual can do better than tell the truth since the outcome is independent of her report. We now show that no other social choice rule is unanimous and strategy-proof. The following proposition has several proofs; the one we present uses [Arrow's impossibility theorem](#).

Proposition 20.2: Gibbard-Satterthwaite theorem

For any **society** $\langle N, X \rangle$ for which X contains at least three alternatives, any **social choice rule** f that is **unanimous** and **strategy-proof** is a dictatorship: for some individual $i^* \in N$, for every profile $(\succsim^i)_{i \in N}$ of preference relations over X we have $f((\succsim^i)_{i \in N}) \succsim^{i^*} x$ for all $x \in X$.

Proof

Let f be a **strategy-proof** and **unanimous social choice rule**.

Step 1 If $f((\succsim^i)_{i \in N}) = x$ and $(\succeq^i)_{i \in N}$ differs from $(\succsim^i)_{i \in N}$ only in that for

some individual j the rank of some alternative y is higher in \succeq^j than it is in \succsim^j , then $f((\succeq^i)_{i \in N}) \in \{x, y\}$.

Proof. Suppose to the contrary that $f((\succeq^i)_{i \in N}) = z$ for some $z \notin \{x, y\}$.

If $z \succ^j x$ then f is not strategy-proof because when j 's preference relation is \succsim^j , if every other individual i reports \succsim^i then j is better off reporting \succeq^j .

If $x \succ^j z$ then also $x \succ^j z$, and f is not strategy-proof because when j 's preference relation is \succeq^j and every other individual i reports $\succsim^i = \succeq^i$ then j is better off reporting \succsim^j (and obtaining the outcome x) than reporting \succeq^j (and obtaining the outcome z). \triangleleft

Step 2 For any two alternatives x and y and any preference profile in which x and y are the top two alternatives for all individuals, the social choice rule f chooses either x or y .

Proof. Assume that there are profiles in which x and y are the top two alternatives for every individual but some other alternative is chosen by f . Let $(\succsim^i)_{i \in N}$ be such a profile with the maximal number of individuals who rank x above y . The maximal number is not $|N|$ since by unanimity if all individuals rank x at the top then f selects x . Let $f((\succsim^i)_{i \in N}) = z$ and let j be an individual for whom $y \succ^j x$. If j reports a preference relation in which x is at the top and y is ranked second then either x or y is chosen by f , and both are better than z , contradicting the strategy-proofness of f . \triangleleft

Step 3 If $(\succsim^i)_{i \in N}$ and $(\succeq^i)_{i \in N}$ are two preference profiles for which x and y are the two top alternatives for every individual and $x \succ^i y$ if and only if $x \succ^i y$ then $f((\succsim^i)_{i \in N}) = f((\succeq^i)_{i \in N})$.

Proof. By Step 2, $f((\succsim^i)_{i \in N}) \in \{x, y\}$. Without loss of generality assume that $f((\succsim^i)_{i \in N}) = x$. We can transform $(\succsim^i)_{i \in N}$ into $(\succeq^i)_{i \in N}$ by a sequence of moves, at each of which we raise one alternative, other than x or y , for one individual while keeping x and y at the top for all individuals. By Step 1 the chosen alternative after each move is either the raised alternative or the alternative chosen by f before the move. By Step 2 the chosen alternative after every move is either x or y . Thus the chosen alternative after each move is x . We conclude that $f((\succeq^i)_{i \in N}) = x$. \triangleleft

Step 4 Given a preference profile $(\succsim^i)_{i \in N}$ define a social binary relation \succsim by $x \succsim y$ if $f((\succsim^i)_{i \in N}) = x$, where for all i the preference relation \succsim^i is obtained from \succsim^i by moving x and y to the top in their original order. The relation \succsim is complete and transitive.

Proof. Completeness follows from Step 2. By Step 3 the relation is anti-symmetric (for no two alternatives $a \succ b$ and $b \succ a$). To verify transitivity, assume that $a \succ b \succ c \succ a$. Consider the profile $(\succeq^i)_{i \in N}$ obtained from $(\succ^i)_{i \in N}$ by moving the three alternatives to the top, preserving their order, in the preference relation of every individual. By an argument analogous to Step 2, $f((\succeq^i)_{i \in N}) \in \{a, b, c\}$. Without loss of generality let $f((\succeq^i)_{i \in N}) = a$. Now, let $(\widetilde{\succ}^i)_{i \in N}$ be the profile obtained from $(\succeq^i)_{i \in N}$ by downgrading b to the third position in all preferences. By Steps 1 and 2, $f((\widetilde{\succ}^i)_{i \in N}) = a$. For each individual the relative order of a and c in the profiles $(\widetilde{\succ}^i)_{i \in N}$ and $(\succ^i)_{i \in N}$ is the same and thus by Step 3 (given $c \succ a$) $f((\widetilde{\succ}^i)_{i \in N}) = c$, a contradiction. \triangleleft

Step 5 *The preference aggregation function F defined in Step 4 is dictatorial.*

Proof. Step 3 implies that F satisfies IIA, and the unanimity of f implies that F satisfies the Pareto property. The existence of a dictator follows from [Arrow's impossibility theorem](#). \triangleleft

Step 6 *There exists an individual i^* such that $f((\succ^i)_{i \in N})$ is \succ^{i^*} -maximal.*

Proof. By Step 5 there is an individual i^* such that $F((\succ^i)_{i \in N}) = \succ^{i^*}$ for any preference profile $(\succ^i)_{i \in N}$. Let $f((\succ^i)_{i \in N}) = x$ and let y be another alternative. By Step 1 we have $f((\succeq^i)_{i \in N}) = x$, where \succeq^i is obtained from \succ^i by moving all alternatives except x and y to below these two alternatives (retaining the order of x and y). By the definition of F at Step 4, $F((\succeq^i)_{i \in N})$ ranks x above y . Since i^* is the dictator $x \succ^{i^*} y$. \triangleleft

Note that the assumption that X contains three alternatives is crucial for this proof. If X contains only two alternatives a and b , then for any K the rule that chooses the alternative a unless K individuals prefer the outcome b is strategy-proof.

Problems

1. *Two alternatives.* Let $X = \{a, b\}$ and assume, as in the body of the chapter, that each individual has strict preferences over the alternatives (she either prefers a to b or b to a). Assume that the number of individuals is odd.

Consider the properties [neutrality](#), [anonymity](#), and [positive responsiveness](#).

- a. For each of the properties give an example of a PAF not satisfying that property but satisfying the other two.
 - b. Show that majority rule is the *unique* PAF satisfying all three properties.
2. *Single-peaked preferences.* Society consists of an odd number of individuals and the set X consists of three political positions. All individuals agree that one of the positions is on the left of the political spectrum, one is in the middle, and one is on the right, so we call the positions L , M , and R . Position M is not at the bottom of any individual's ranking, so that no individual has the preference $L \succ R \succ M$ or the preference $R \succ L \succ M$.

Show that for any preference profile in this restricted domain, majority rule induces a preference relation over X .

3. A group of n individuals has to choose an alternative from a set X . Each individual has a favorite alternative. A decision method is a function F that attaches a member of X to each profile (x_1, \dots, x_n) of favorite alternatives.
 - a. Define formally the property of neutrality, which requires that a decision method treats all alternatives equally. Give two different examples of decision methods that satisfy this property.
 - b. Define formally the Pareto property, which requires that for any $x \in X$ a decision method chooses x if all individuals choose x . Give an example of a decision method that does not satisfy this property.
 - c. Define formally the notion of a dictatorial decision method.

Say that a decision method F satisfies property I if whenever (x_1, \dots, x_n) and (y_1, \dots, y_n) satisfy the condition

- for some a the set of individuals who choose a in (x_1, \dots, x_n) is equal to the set of individuals who choose a in (y_1, \dots, y_n)

then $F(x_1, \dots, x_n) = a$ if and only if $F(y_1, \dots, y_n) = a$.

For example, $F(a, b, c, a) = a$ if and only if $F(a, c, d, a) = a$.

- d. Show that if X contains only two alternatives and n is odd then there exists a decision method that is neutral, satisfies the Pareto property and property I , and is not dictatorial.
 - e. Assume that X contains n elements and that $n \geq 3$. Show that the only methods that are neutral and satisfy property I are dictatorial.

4. *Classification.* A group N of individuals discusses a group X of objects. Each individual is associated with a partition of the objects into classes. For example, if $X = \{a, b, c\}$ then each individual is associated with one of the five possible partitions (i) each object is in a distinct class, (ii) all objects are in one class, and (iii) one of the objects (a , b , or c) is in its own class and the other two are in the same class. An aggregation method attaches one partition of X to each profile of partitions of X .
- Show that the rule that puts two objects in the same class if a majority of the individuals' partitions put them in the same class is not an aggregation method.
 - Show that the rule that puts two objects in the same class if the partitions of all members of a certain group $G \subseteq N$ put them in the same class is an aggregation method.
5. *Ranking participants in a tournament.* A group of n players compete in a round-robin tournament. Each match ends with one of the players winning the match (a tie is not possible). A *ranking method* attaches to each possible set of results a ranking of the players (possibly with indifferences).

Consider the following three properties of ranking methods.

Anonymity The method treats all players equally.

Monotonicity Let R_1 be a set of results for which the method ranks player i at least as high as player j . Let R_2 be a set of results identical to R_1 except that some player wins against i in R_1 and loses against i in R_2 . Then i is ranked (strictly) above j in R_2 .

Independence The relative ranking of any two players is independent of the results for the matches in which they do not participate.

- A familiar method ranks players by their number of victories. Verify that this method satisfies the three properties.
 - For each of the three properties give an example of a method that does not satisfy the property although it satisfies the other two.
6. *Median.* When the space of preferences is restricted, there exist nondictatorial strategy-proof social choice rules. Assume that the number of individuals is $2m + 1$ (for some positive integer m), $X = [0, 1]$, and each individual i has single-peaked preferences (there is $z^i \in X$ such that if $z^i < a < b$ or $b < a < z^i$ then $z^i \succ^i a \succ^i b$). Show that the social choice rule that attaches to every preference profile the median of the peaks is strategy-proof.

Notes

Proposition 20.1 is due to Arrow (1951). The proof we give is due to Geanakoplos (2005). Proposition 20.2 appears in Gibbard (1973) and Satterthwaite (1975). Problem 1 is based on May (1952) and Problem 5 is based on Rubinstein (1980).

