Social Choice 11

Social choice theory studies the aggregation of individual preferences into a common or social preference. It overlaps with several social science disciplines, such as political theory (e.g., voting for Parliament, or for a president) and game theory (e.g., voters may vote strategically, or candidates may choose positions strategically).

In the classical model of social choice, there is a finite number of agents who have preferences over a finite number of alternatives. These preferences are either aggregated into a social preference according to a so-called social welfare function, or result in a common alternative according to a so-called social choice function.

The main purpose of this chapter is to review two classical results, namely Arrow's Theorem and the Gibbard-Satterthwaite Theorem. The first theorem applies to social welfare functions and says that, if the social preference between any two alternatives should only depend on the individual preferences between these alternatives and, thus, not on individual preferences involving other alternatives, then basically the social welfare function must be dictatorial. The second theorem applies to social choice functions and says that, basically, the only social choice functions that are invulnerable to strategic manipulation are the dictatorial ones. These results are often referred to as 'impossibility theorems' since dictatorships are generally regarded undesirable.

Section 11.1 is introductory. Section 11.2 discusses Arrow's Theorem and Sect. 11.3 the Gibbard-Satterthwaite Theorem.

11.1 Introduction and Preliminaries

11.1.1 An Example

Suppose there are three agents (individuals, voters) who have strict preferences over a set of five alternatives $(a_1, \ldots, a_5)$, as given in Table 11.1.

H. Peters, *Game Theory*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-662-46950-7_11

Table 11.1 Borda scores

Agent	a_1	a_2	a_3	a_4	a_5
1	5	1	3	2	4
2	1	2	3	4	5
3	3	4	5	2	1

In this table the preferences of the players are represented by the *Borda scores*: the best alternative of an agent obtains 5 points, the second best 4 points, etc., until the worst alternative which obtains 1 point. For instance, agent 1 has the preference $a_1P_1a_5P_1a_3P_1a_4P_1a_2$ in the notation to be introduced below. We use the Borda scores as a convenient way to represent these preferences and, more importantly, to obtain an example of a social welfare as well as a social choice function.

First, suppose that we want to extract a common social ranking of the alternatives from the individual preferences. One way to do this is to add the Borda scores per alternative. In the example this results in $9, 7, 11, 8, 10$ for a_1, a_2, a_3, a_4, a_5, respectively, resulting in the social ranking $a_3Pa_5Pa_1Pa_4Pa_2$. If we just want to single out one alternative, then we could take the one with the maximal Borda score, in this case alternative a_3. In the terminology to be introduced formally below, Borda scores give rise to a social welfare as well as a social choice function.[1]

One potential drawback of using Borda scores to obtain a social ranking is, that the ranking between two alternatives may not just depend on the individual preferences between these two alternatives. For instance, suppose that agent 1's preference would change to $a_1P_1a_4P_1a_5P_1a_3P_1a_2$. Then the Borda scores would change to $9, 7, 10, 10, 9$ for a_1, a_2, a_3, a_4, a_5, respectively, resulting in the social ranking $a_3Ia_4Pa_1Ia_5Pa_2$ (where I denotes indifference). Observe that no agent's preference between a_1 and a_4 has changed, but that socially this preference is reversed. This is not a peculiarity of using Borda scores: Arrow's Theorem, to be discussed in Sect. 11.2, states that under some reasonable additional assumptions the only way to avoid this kind of preference reversal is to make one agent the dictator, i.e., to have the social preference coincide with the preference of one fixed agent.

Another potential drawback of using the Borda scores in order to single out a unique alternative is that this method is vulnerable to strategic manipulation. For instance, suppose that agent 1 would lie about his true preference given in Table 11.1 and claim that his preference is $a_1P_1a_5P_1a_2P_1a_4P_1a_3$ instead. Then the Borda scores would change to $9, 9, 9, 8, 10$ for a_1, a_2, a_3, a_4, a_5, respectively, resulting in the chosen alternative a_5 instead of a_3. Since agent 1 prefers a_5 over a_3 according to his *true* preference, he gains by this strategic manipulation. Again, this phenomenon

[1]Ties may occur, but this need not bother us here.

is not a peculiarity of the Borda method: the Gibbard-Satterthwaite Theorem in Sect. 11.3 shows that again under some reasonable additional assumptions the only way to avoid it is to make one fixed agent a dictator.

11.1.2 Preliminaries

Let $A = \{a_1, \ldots, a_m\}$ be the set of *alternatives*. To keep things interesting we assume $m \geq 3$.[2] The set of *agents* is denoted by $N = \{1, \ldots, n\}$. We assume $n \geq 2$.

A *binary relation* on A is a subset of $A \times A$. In our context, for a binary relation R on A we usually write aRb instead of $(a, b) \in R$ and interpret this as an agent or society (weakly) preferring a over b. Well-known conditions for a binary relation R on A are:

(a) *Reflexivity*: aRa for all $a \in A$.
(b) *Completeness*: aRb or bRa for all $a, b \in A$ with $a \neq b$.
(c) *Antisymmetry*: For all $a, b \in A$, if aRb and bRa, then $a = b$.
(d) *Transitivity*: For all $a, b, c \in A$, aRb and bRc imply aRc.

A *preference* on A is a reflexive, complete and transitive binary relation on A. For a preference R on A we write aPb if aRb and not bRa; and aIb if aRb and bRa. The binary relations P and I are called the *asymmetric* and *symmetric parts* of R, respectively, and are interpreted as strict preference and indifference. Check (Problem 11.1) that P is antisymmetric and transitive but not reflexive and not necessarily complete, and that I is reflexive and transitive but not necessarily antisymmetric and not necessarily complete. By $\mathcal{L}^*$ we denote the set of all preferences on A, and by $\mathcal{L} \subseteq \mathcal{L}^*$ the set of all antisymmetric (i.e., *strict*) preferences on A. In plain words, elements of $\mathcal{L}^*$ order the elements of A but allow for indifferences, while elements of $\mathcal{L}$ order the elements of A strictly.[3]

In what follows, it is assumed that agents have strict preferences while social preferences may have indifferences. A strict preference *profile* is a list $(R_1, \ldots, R_i, \ldots, R_n)$, where R_i is the strict preference of agent i. Hence, $\mathcal{L}^N$ denotes the set of all strict preference profiles. A *social choice function* is a map $f : \mathcal{L}^N \to A$, i.e., it assigns a unique alternative to every profile of strict preferences. A *social welfare function* is a map $F : \mathcal{L}^N \to \mathcal{L}^*$, i.e., it assigns a (possibly non-strict) preference to every profile of strict preferences.

[2] See Problem 11.7 for the case $m = 2$.

[3] Elements of $\mathcal{L}$ are usually called *linear orders* and those of $\mathcal{L}^*$ *weak orders*.

11.2 Arrow's Theorem

In this section the focus is on social welfare functions. We formulate three properties for a social welfare function $F : \mathcal{L}^N \to \mathcal{L}^*$. We call F:

(a) *Pareto Efficient* (PE) if for each profile $(R_1, \ldots, R_n) \in \mathcal{L}^N$ and all $a, b \in A$, if $a \neq b$ and aR_ib for all $i \in N$, then aPb, where P is the asymmetric part of $R = F(R_1, \ldots, R_n)$.
(b) *Independent of Irrelevant Alternatives* (IIA) if for all $(R_1, \ldots, R_n) \in \mathcal{L}^N$ and $(R'_1, \ldots, R'_n) \in \mathcal{L}^N$ and all $a, b \in A$, if $aR_ib \Leftrightarrow aR'_ib$ for all $i \in N$, then $aRb \Leftrightarrow aR'b$, where $R = F(R_1, \ldots, R_n)$ and $R' = F(R'_1, \ldots, R'_n)$.
(c) *Dictatorial* (D) if there is an $i \in N$ such that $F(R_1, \ldots, R_n) = R_i$ for all $(R_1, \ldots, R_n) \in \mathcal{L}^N$.

Pareto Efficiency requires that, if all agents prefer an alternative a over an alternative b, then the social ranking should also put a above b. Independence of Irrelevant Alternatives says that the social preference between two alternatives should only depend on the agents' preferences between these two alternatives and not on the position of any other alternative.[4] Dictatoriality says that the social ranking is always equal to the preference of a fixed agent, the *dictator*. Clearly, there are exactly n dictatorial social welfare functions.

The first two conditions are usually regarded as desirable but the third clearly not. Unfortunately, Arrow's Theorem implies that the first two conditions imply the third one.[5]

Theorem 11.1 (Arrow's Theorem) *Let F be a Pareto Efficient and IIA social welfare function. Then F is dictatorial.*

Proof

Step 1 Consider a profile in $\mathcal{L}^N$ and two distinct alternatives $a, b \in A$ such that every agent ranks a on top and b at bottom. By Pareto Efficiency, the social ranking assigned by F must also rank a on top and b at bottom.

Now change agent 1's ranking by raising b in it one position at a time. By IIA, a is ranked socially (by F) on top as long as b is still below a in the preference of agent 1. In the end, if agent 1 ranks b first and a second, we have a or b on top of the social ranking by Pareto efficiency of F. If a is still on top in the social ranking, then continue the same process with agents 2, 3, etc., until we reach some agent k such that b is on top of the social ranking after moving b above a in agent k's preference.

[4] Although there is some similarity in spirit, this condition is not in any formal sense related to the IIA condition in bargaining, see Sect. 10.1 or Chap. 21.

[5] For this reason the theorem is often referred to as Arrow's Impossibility Theorem.

Table 11.2 Step 1 of the proof of Theorem 11.1, agent k ranks a above b

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
b	$\cdots$	b	a	a	$\cdots$	a	a	a
a	$\cdots$	a	b	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	b	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	b	$\cdots$	b	$\cdot$	

Table 11.3 Step 1 of the proof of Theorem 11.1, agent k ranks b above a

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
b	$\cdots$	b	b	a	$\cdots$	a	b	b
a	$\cdots$	a	a	$\cdot$	$\cdots$	$\cdot$	a	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	b	$\cdots$	b	$\cdot$	

Table 11.4 Step 2 of the proof of Theorem 11.1, arising from Table 11.2

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
b	$\cdots$	b	a	$\cdot$	$\cdots$	$\cdot$	a	a
$\cdot$	$\cdots$	$\cdot$	b	$\cdot$	$\cdots$	$\cdot$	b	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	a	$\cdots$	a	$\cdot$	
a	$\cdots$	a	$\cdot$	b	$\cdots$	b	$\cdot$	

Table 11.5 Step 2 of the proof of Theorem 11.1, arising from Table 11.3

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
b	$\cdots$	b	b	$\cdot$	$\cdots$	$\cdot$	b	b
$\cdot$	$\cdots$	$\cdot$	a	$\cdot$	$\cdots$	$\cdot$	$\cdot$	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	a	
$\cdot$	$\cdots$	$\cdot$	$\cdot$	a	$\cdots$	a	$\cdot$	
a	$\cdots$	a	$\cdot$	b	$\cdots$	b	$\cdot$	

Tables 11.2 and 11.3 give the situations just before and just after b is placed above a in agent k's preference.[6]

Step 2 Now consider Tables 11.4 and 11.5.
The profile in Table 11.4 arises from the one in Table 11.2 by moving a to the last position for agents $i < k$ and to the second last position for agents $i > k$. In exactly the same way, the profile in Table 11.5 arises from the one in Table 11.3. Then IIA applied to Tables 11.3 and 11.5 implies that b is socially top-ranked in Table 11.5. Next, IIA applied to the transition from Table 11.5 to Table 11.4 implies that in

[6] In these tables and also the ones below, we generically denote all preferences by $R_1, \ldots, R_n$. The last column in every table will be used in Sect. 11.3.

Table 11.6 Step 3 of the proof of Theorem 11.1

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
·	$\cdots$	·	a	·	$\cdots$	·	a	a
·	$\cdots$	·	c	·	$\cdots$	·	·	
·	$\cdots$	·	b	·	$\cdots$	·	·	
c	$\cdots$	c	·	c	$\cdots$	c	·	
b	$\cdots$	b	·	a	$\cdots$	a	·	
a	$\cdots$	a	·	b	$\cdots$	b	·	

Table 11.7 Steps 4 and 5 of the proof of Theorem 11.1

R_1	$\cdots$	R_{k-1}	R_k	R_{k+1}	$\cdots$	R_n	F	f
·	$\cdots$	·	a	·	$\cdots$	·	a	a
·	$\cdots$	·	c	·	$\cdots$	·	·	
·	$\cdots$	·	b	·	$\cdots$	·	c	
c	$\cdots$	c	·	c	$\cdots$	c	·	
b	$\cdots$	b	·	b	$\cdots$	b	b	
a	$\cdots$	a	·	a	$\cdots$	a	·	

Table 11.4 b must still be socially ranked above every alternative except perhaps a. But IIA applied to the transition from Table 11.2 to Table 11.4 implies that in Table 11.4 a must still be socially ranked above every alternative. This proves that the social rankings in Tables 11.4 and 11.5 are correct.

Step 3 Consider a third alternative c distinct from a and b. The social ranking in Table 11.6 is obtained by from Table 11.4 by applying IIA.

Step 4 Consider the profile in Table 11.7, obtained from the profile in Table 11.6 by switching a and b for agents $i > k$. By IIA applied to the transition from Table 11.6 to Table 11.7, we have that a must still be socially ranked above every alternative except possibly b. However, b must be ranked below c by Pareto efficiency, which shows that the social ranking in Table 11.7 is correct.

Step 5 Consider any arbitrary profile in which agent k prefers a to b. Change the profile by moving c between a and b for agent k and to the top of every other agent's preference (if this is not already the case). By IIA this does not affect the social ranking of a vs. b. Since the preference of every agent concerning a and c is now as in Table 11.7, IIA implies that a is socially ranked above c, which itself is socially ranked above b by Pareto Efficiency. Hence, by transitivity of the social ranking we may conclude that a is socially ranked above b whenever it is preferred by agent k over b. By repeating the argument with the roles of b and c reversed, and recalling that c was an arbitrary alternative distinct from a and b, we may conclude that the social ranking of a is above some alternative whenever agent k prefers a to that alternative: k is a 'dictator' for a. Since a was arbitrary, we can repeat the whole argument to conclude that there must be a dictator for every alternative. Since there cannot be distinct dictators for distinct alternatives, there must be a single dictator for all alternatives. ■

11.3 The Gibbard-Satterthwaite Theorem

The Gibbard-Satterthwaite Theorem applies to social choice functions. We start by listing the following possible properties of a social choice function $f : \mathcal{L}^N \to A$. Call f:

(a) *Unanimous* (UN) if for each profile $(R_1, \ldots, R_n) \in \mathcal{L}^N$ and each $a \in A$, if aR_ib for all $i \in N$ and all $b \in A \setminus \{a\}$, then $f(R_1, \ldots, R_n) = a$.
(b) *Monotonic* (MON) if for all profiles $(R_1, \ldots, R_n) \in \mathcal{L}^N$ and $(R'_1, \ldots, R'_n) \in \mathcal{L}^N$ and all $a \in A$, if $f(R_1, \ldots, R_n) = a$ and $aR_ib \Rightarrow aR'_ib$ for all $b \in A \setminus \{a\}$ and $i \in N$, then $f(R'_1, \ldots, R'_n) = a$.
(c) *Dictatorial* (D) if there is an $i \in N$ such that $f(R_1, \ldots, R_n) = a$ where aR_ib for all $b \in A \setminus \{a\}$, for all $(R_1, \ldots, R_n) \in \mathcal{L}^N$.
(d) *Strategy-Proof* (SP) if for all profiles $(R_1, \ldots, R_n) \in \mathcal{L}^N$ and $(R'_1, \ldots, R'_n) \in \mathcal{L}^N$ and all $i \in N$, if $R'_j = R_j$ for all $j \in N \setminus \{i\}$, then $f(R_1, \ldots, R_n)\ R_i\ f(R'_1, \ldots, R'_n)$.

Unanimity requires that, if all agents have the same top alternative, then this alternative should be chosen. Monotonicity says that, if some alternative a is chosen and the profile changes in such a way that a is still preferred by every agent over all alternatives over which it was originally preferred, then a should remain to be chosen.[7] Dictatoriality means that there is a fixed agent whose top element is always chosen. Strategy-Proofness says that no agent can obtain a better chosen alternative by lying about his true preference.

In accordance with mathematical parlance, call a social choice function $f : \mathcal{L}^N \to A$ *surjective* if for every $a \in A$ there is some profile $(R_1, \ldots, R_n) \in \mathcal{L}^N$ such that $f(R_1, \ldots, R_n) = a$. Hence, each a is chosen at least once.[8] The Gibbard-Satterthwaite Theorem is as follows.

Theorem 11.2 (Gibbard-Satterthwaite Theorem) *Let $f : \mathcal{L}^N \to A$ be a surjective and strategy-proof social choice function. Then f is dictatorial.*

Since surjectivity is implied by unanimity, Theorem 11.2 also holds with unanimity instead of surjectivity.

We will prove the Gibbard-Satterthwaite Theorem by using the next theorem, which is a variant of the Muller-Satterthwaite Theorem.

Theorem 11.3 (Muller-Satterthwaite) *Let $f : \mathcal{L}^N \to A$ be a unanimous and monotonic social choice function. Then f is dictatorial.*

Proof of Theorem 11.2 We prove that f is unanimous and monotonic. The result then follows from Theorem 11.3.

[7] This property is also called *Maskin Monotonicity*, after Maskin (1999).

[8] In the social choice literature this property is sometimes called *citizen-sovereignty*.

Suppose that $f(R_1, \dots, R_n) = a$ for some profile $(R_1, \dots, R_n) \in \mathcal{L}^N$ and some alternative $a \in A$. Let $i \in N$ and let $(R'_1, \dots, R'_n) \in \mathcal{L}^N$ be a profile such that for all $j \in N \setminus \{i\}$ we have $R'_j = R_j$ and for all $b \in A \setminus \{a\}$ we have aR'_ib if aR_ib. We wish to show that $f(R'_1, \dots, R'_n) = a$. Suppose, to the contrary, that $f(R'_1, \dots, R'_n) = b \neq a$. Then SP implies aR_ib, and hence aR'_ib. Again by SP, however, bR'_ia, hence, by antisymmetry of R'_i, $a = b$, a contradiction. This proves $f(R'_1, \dots, R'_n) = a$.

Now suppose that $(R'_1, \dots, R'_n) \in \mathcal{L}^N$ is a profile such that for all $i \in N$ and all $b \in A \setminus \{a\}$ we have aR'_ib if aR_ib. By applying the argument in the preceding paragraph n times, it follows that $f(R'_1, \dots, R'_n) = a$. Hence, f is monotonic.

To prove unanimity, suppose that $(R_1, \dots, R_n) \in \mathcal{L}^N$ and $a \in A$ such that aR_ib for all $i \in N$ and $b \in A \setminus \{a\}$. By surjectivity there is $(R'_1, \dots, R'_n) \in \mathcal{L}^N$ with $f(R'_1, \dots, R'_n) = a$. By monotonicity we may move a to the top of each agent's preference and still have a chosen. Next, again by monotonicity, we may change each agent i's preference to R_i without changing the chosen alternative, i.e., $f(R_1, \dots, R_n) = a$. Hence, f is unanimous. ■

Proof of Theorem 11.3 The proof parallels the proof of Theorem 11.1 and uses analogous steps and the same tables.

Step 1 Consider a profile in $\mathcal{L}^N$ and two distinct alternatives $a, b \in A$ such that every agent ranks a on top and b at bottom. By unanimity, f chooses a.
Now change agent 1's ranking by raising b in it one position at a time. By MON, a is chosen by f as long as b is still below a in the preference of agent 1. In the end, if agent 1 ranks b first and a second, we have a or b chosen by f, again by MON. If a is still chosen, then continue the same process with agents 2, 3, etc., until we reach some agent k such that b is chosen after moving b above a in agent k's preference. Tables 11.2 and 11.3 give the situations just before and just after b is placed above a in agent k's preference.

Step 2 Now consider Tables 11.4 and 11.5. The profile in Table 11.4 arises from the one in Table 11.2 by moving a to the last position for agents $i < k$ and to the second last position for agents $i > k$. In exactly the same way, the profile in Table 11.5 arises from the one in Table 11.3.
Then MON applied to Tables 11.3 and 11.5 implies that b is chosen in Table 11.5. Next, MON applied to the transition from Table 11.5 to Table 11.4 implies that in Table 11.4 the choice must be either b or a. Suppose b would be chosen. Then MON applied to the transition from Table 11.4 to Table 11.2 implies that in Table 11.2 b must be chosen as well, a contradiction. Hence, a is chosen in Table 11.4. This proves that the choices by f in Tables 11.4 and 11.5 are correct.

Step 3 Consider a third alternative c distinct from a and b. The choice in Table 11.6 is obtained by from Table 11.4 by applying MON.

Step 4 Consider the profile in Table 11.7, obtained from the profile in Table 11.6 by switching a and b for agents $i > k$. If the choice in Table 11.7 were some d unequal to a or b, then by MON it would also be d in Table 11.6, a contradiction. If it were b, then by MON it would remain b even if c would be moved to the top of every agent's preference, contradicting unanimity. Hence, it must be a.

Step 5 Consider any arbitrary profile with a at the top of agent k's preference. Such a profile can always be obtained from the profile in Table 11.7 without worsening the position of a with respect to any other alternative in any agent's preference. By MON therefore, a must be chosen whenever it is at the top of agent k's preference, so k is a 'dictator' for a. Since a was arbitrary, we can find a dictator for every other alternative but, clearly, these must be one and the same agent. Hence, this agent is the dictator. ■

There is a large literature that tries to escape the negative conclusions of Theorems 11.1–11.3 by adapting the model and/or restricting the domain. Examples are provided in Problems 6.23 and 6.24.

11.4 Problems

11.1. *Preferences*
Let R be a preference on A, with symmetric part I and asymmetric part P.

(a) Prove that P is antisymmetric and transitive but not reflexive and not necessarily complete.
(b) Prove that I is reflexive and transitive but not necessarily complete and not necessarily antisymmetric.

11.2. *Pairwise Comparison*
For a profile $r = (R_1, \ldots, R_n) \in \mathcal{L}^N$ and $a, b \in A$ define

$$N(a, b, r) = \{i \in N \mid aR_ib\} \ ,$$

i.e., $N(a, b, r)$ is the set of agents who (strictly) prefer a to b in the profile r. With r we can associate a binary relation $C(r)$ on A by defining $aC(r)b :\Leftrightarrow |N(a, b, r)| \geq |N(b, a, r)|$ for all $a, b \in A$. If $aC(r)b$ we say that 'a beats b by pairwise majority'.

(a) Is $C(r)$ reflexive? Complete? Antisymmetric?
(b) Show that $C(r)$ is not transitive, by considering the famous Condorcet profile for $N = \{1, 2, 3\}$ and $A = \{a, b, c\}$: aR_1bR_1c, bR_2cR_2a, cR_3aR_3b.
(c) Call a a *Condorcet winner* if $|N(a, b, r)| > |N(b, a, r)|$ for all $b \in A \setminus \{a\}$. Is there a Condorcet winner in the example in Sect. 11.1?

11.3. *Independence of the Conditions in Theorem 11.1*
Show that the conditions in Theorem 11.1 are independent. That is, exhibit a social welfare function that is Pareto efficient and does not satisfy IIA or dictatoriality, and one that satisfies IIA and is not dicatorial nor Pareto efficient.

11.4. *Independence of the Conditions in Theorem 11.2*
Show that the conditions in Theorem 11.2 are independent.

11.5. *Independence of the Conditions in Theorem 11.3*
Show that the conditions in Theorem 11.3 are independent.

11.6. *Copeland Score and Kramer Score*
The *Copeland score* of an alternative $a \in A$ at a profile $r = (R_1, \ldots, R_n) \in \mathcal{L}^N$ is defined by

$$c(a, r) = |\{b \in A \mid N(a, b, r) \geq N(b, a, r)\}| ,$$

i.e., the number of alternatives that a beats (cf. Problem 11.2). The *Copeland ranking* is obtained by ranking the alternatives according to their Copeland scores.

(a) Is the Copeland ranking a preference? Is it antisymmetric? Does the derived social welfare function satisfy IIA? Pareto efficiency?

The *Kramer score* of an alternative $a \in A$ at a profile $r = (R_1, \ldots, R_n) \in \mathcal{L}^N$ is defined by

$$k(a, r) = \min_{b \in A \setminus \{a\}} |N(a, b, r)| ,$$

i.e., the worst score among all pairwise comparisons. The *Kramer ranking* is obtained by ranking the alternatives according to their Kramer scores.

(b) Is the Kramer ranking a preference? Is it antisymmetric? Does the derived social welfare function satisfy IIA? Pareto efficiency?

11.7. *Two Alternatives*
Show that Theorems 11.1–11.3 no longer hold if there are just two alternatives, i.e., if $m = 2$.

11.5 Notes

For a general overview of social choice theory see Arrow et al. (2002, 2011). For Arrow's Theorem see Arrow (1963). For the Gibbard-Satterthwaite Theorem see Gibbard (1973) and Satterthwaite (1975). Both these theorems are closely related:

indeed, the proof of the Gibbard-Satterthwaite Theorem in Gibbard (1973) uses Arrow's Theorem. The presentation in this chapter closely follows that in Reny (2001), which is both simple and elegant, and which shows the close relation between the two results.

For the Borda scores and the so-called Borda rule see de Borda (1781). The Muller-Satterthwaite Theorem is from Muller and Satterthwaite (1977).

De Condorcet (1785) was the first to explicitly discuss the notion of a Condorcet winner; see Gehrlein (2006) for a comprehensive study of the so-called *Condorcet paradox* in Problem 11.2(b).

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: Wiley.

Arrow, K. J., Sen, A. K., & Suzumura, K. (Eds.), (2002). *Handbook of social choice and welfare* (Vol. 1). Amsterdam: North-Holland.

Arrow, K. J., Sen, A. K., & Suzumura, K. (Eds.), (2011). *Handbook of social choice and welfare* (Vol. 2). Amsterdam: North-Holland.

de Borda, J. C. (1781). *Mémoires sur les élections au scrutin*. Paris: Histoire de l'Académie Royale des Sciences.

de Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie Royale.

Gehrlein, W. V. (2006). *Condorcet's paradox*. Berlin: Springer.

Gibbard, A. (1973). Manipulation of voting schemes: a general result. *Econometrica, 41*, 587–601.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies, 66*, 23–38.

Muller, E., & Satterthwaite, M. A. (1977). The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory, 14*, 412–418.

Reny, P. J. (2001). Arrow's theorem and the Gibbard-Satterthwaite theorem: a unified approach. *Economics Letters, 70*, 99–105.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory, 10*, 187–217.