4 Voting and Social Choice

4.1 Ordinal Welfarism

As discussed in section 1.3, ordinal welfarism pursues the welfarist program in those situations where the cardinal measurement of individual welfares is either unfeasible, unreliable, or ethically untenable.

Consider voting among several candidates competing for a certain political office. Who would think of measuring the relative impact of electing Jones or Smith on the individual welfare of each and every citizen? Say that Ann favors Jones and Bob favors Smith. It is typically impossible to decide whether Ann's utility gain from having Jones in office rather than Smith is larger than Bob's gain from having Smith rather than Jones. Such an evaluation would require that each voter forms expectations about all decisions influenced by the election of Jones or Smith and computes his or her expected utility. The complexity of these computations destroys the plausibility of the cardinal utility model: electors simply do not perform elaborate computations of cardinal welfare, and even the task of ranking all candidates from best to worst is challenging, given limited information about the consequences of the election.

In most real life elections, in the large or in the small, voters are not asked to express more than an "ordinal" opinion, namely a preference ordering (which may or may not allow for indifferences) of the names on the ballot. There is more to this restriction than the practicality and simplicity of an ordinal message.

Recall from section 3.1 the strong ethical objections to the interpersonal comparison of cardinal welfares. If the outcome of the election depends on the intensity of the voters' feelings and emotions about the candidates, a minority of fanatics will influence the outcome more than the quiet majority; worse yet, a subset of voters *faking* fanaticism are more influential than the honest, truth-telling voters.

The central postulate of ordinal welfarism is that individual welfare is entirely captured by a preference ordering of the possible outcomes, also called states of the world. This ordering is a binary relation R of the set A of possible outcomes (A is often called the choice set). The relation xRy reads "x is at least as good as y," or "welfare at x is not below welfare at y." The relation R is assumed to be *complete* and *transitive*. Completeness of R means that all pairs x, y in A can be compared: at least one of xRy and yRx holds. If xRy but not yRx, we say that x is strictly preferred to y and write xPy; if xRy and yRx, we say that the choice between x and y is a matter of indifference (or, abusing language, that x and y are "indifferent") and we write xIy. Transitivity of R means that xRy and yRz imply xRz. In particular, strict preferences are transitive (xPy and yPz imply xPz), and so are indifferences (xIy and yIz imply xIz).

The preference relation R expresses the opinion, tastes, or values of a certain agent over the outcomes in A, and it pointedly avoids any statement about the intensity of these preferences. The empirical basis of R is the *choice* between outcomes: if an agent with preferences R is presented with a choice between two states of the world x and y, she will choose x if xPy and y if yPx. If presented with a choice over a subset $\{x, y, z, \ldots\}$ of outcomes, she will pick one of the outcomes ranked highest in R, namely a^* such that a^*Rx, a^*Ry, a^*Rz, and so on.

The choices made by a given agent over the various subsets of outcomes that she may be presented with, are entirely determined by her preference relation R. Conversely, if we observe a series of choices, we may or may not be able to identify an underlying preference relation *rationalizing* these choices, in the sense that the observed choices are the highest ranked for R among the feasible choices. For instance, imagine that Ann has been presented with three successive pairs $\{x, y\}$, $\{y, z\}$, and her choices are as follows: choose x from $\{x, y\}$, choose y from $\{y, z\}$, or choose z from $\{x, z\}$. If the preference relation R rationalizes these choices, it must have xPy because x was unambiguously chosen from x, y (Ann could say "both choices are okay" but not choose to do so), and similarly we must have yPz and zPx, in contradiction of the transitivity of R. Another configuration ruling out the existence of R is

$$\begin{aligned} &\text{choose } x \text{ from } \{x, y\} \\ &\text{choose } y \text{ from } \{x, y, z\} \end{aligned} \tag{1}$$

Here the first choice reveals xPy, but the second implies yPx, yPz, in contradiction of the definitions of P.

In the configuration (1) the peculiarity of the observed choices is illustrated by the following classic air travel story. The flight attendant asks this passenger if he will have fish or chicken for his meal, and his answer is "fish." A minute later he comes back announcing that the pasta meal is still available, and would he like to change his mind? Yes, says the passenger, in that case I will have chicken.

The choices over all conceivable feasible subsets of outcomes are said to be *rational* (or rationalizable) if there exists a preference relation R such that for any feasible subset B of outcomes, the outcomes (or outcomes) $S(B)$ selected from B are precisely the highest ranked outcomes in B according to $R : xPy$ if x is in $S(B)$ and y is not in $S(B)$; xIx' if x and x' are both in $S(B)$. The formal theory of rational choice explores in great detail what restrictions on the various choices over the various subsets guarantee that these choices can be rationalized by a preference relation.

The identification of welfare with preferences, and of preferences with choice, is an intellectual construction at the center of modern economic thinking. We will refer to this construction as the ordinal approach. It eschews heroic assumptions about cardinal measurement of utility, and offers a testable empirical basis for the construction of individual preferences, namely the actual choices made by the agents. Social choice theory adapts the welfarist program—namely the definition of just compromises between the conflicting goals of maximizing different individual welfares—to the ordinal approach.

The concept of Pareto optimality, already central in chapter 3, is clearly an ordinal concept. Outcome y is Pareto superior to outcome x if $y R_j x$ for all agent j and $y P_i x$ for at least one agent i: everyone is at least as satisfied with y as with x and at least one agent is strictly better off with y.

The main new feature of the ordinal context is that individual welfare can no longer be separated from the set A of outcomes to which it applies: the binary relation R bears on A and cannot be defined in a vacuum. Contrast this with the cardinal approach, where we can compare utility levels $u_1, u_1', u_2, u_2', \ldots$ without specifying the outcomes from which these utilities are derived. Therefore, in the ordinal world, collective decision-making can only be defined if we specify the set A of feasible outcomes (states of the world), and for each agent i a preference relation R_i on A. The focus is on the distribution of decision power, namely how the configuration of these relations R_i—called the profile of preferences—affects the choice of an outcome in A, or affects the collective preference relation on A—in the cases of voting and preference aggregation, respectively—as explained below.

We define now the two central models of social choice theory. A *voting* problem specifies the set A of outcomes, the set N of agents, and a profile of preferences; the problem is to "elect" an outcome a in A from these data. A systematic solution of this problem (a rule for selecting an outcome from any profile of preferences) is a pure example of a public contract in the sense of section 1.5. Key to the discussion starting in the next section is the knowledge of which subsets T of N "control" the choice of a in the sense that if all voters in T have a as their first choice, a will be elected. For instance, under majority voting, any coalition T containing strictly more agents than N/T controls the outcome of the election.

The voting model has two complementary interpretations, normative and strategic. In the normative one, the benevolent dictator discovers the profile of preferences R_i and enforces the compromise outcome deemed just by the voting method. Alternatively, the preference relation R_i is private information to agent i and a voting rule is a decentralized decision process enforced by the public authority. Every voter reports a preference relation $\tilde{R}_i$ to the central agency, who takes these messages on face value to compute the winning outcome. As the agency has no way to determine whether the reported preference relation $\tilde{R}_i$ is indeed the true relation R_i, each agent is free to report a nontruthful relation if this serves his or her interest better than reporting the truth. Thus, in the second interpretation of the voting model, the issue is strategic voting, and how the central agency can avoid misreporting.

The second model of social choice theory is the *preference aggregation* problem. Here we associate to a preference profile a collective preference. Much as in chapter 3 a collective utility function associates a cardinal index to any profile of cardinal utilities; the aggregation method computes an ordinal preference relation from any profile of preference relations. The anthropomorphic bias is the same in both cases, but the new ordinal model is technically more involved.

We are given, as in the voting model, the outcome set A, agent set N, and a profile of preferences, but we are now looking for a collective preference R on A, one that will order all outcomes instead of selecting just one "best" outcome. The two problems are closely related, and under the axiom independence of irrelevant alternatives (section 4.6), they are formally equivalent.

The voting and preference aggregation problems are truly the most general microeconomic models of collective decision-making because they make no restrictive assumption, neither on the set A of outcomes or on the admissible preference profiles of the agents. Therefore these models encompass, in principle, all resource allocation problems from chapter 5 onward. On the other hand, the extreme generality of the model leads to two severe impossibility results, namely Arrow's theorem about preference aggregation (section 4.6), and Gibbard-Satterthwaite's theorem about strategic voting (briefly discussed at the end of section 4.4). More palatable results obtain when the domain of individual preferences is suitably restricted, and two important examples of such restrictions are discussed in sections 4.4 and 4.5.

4.2 Condorcet versus Borda

The two most important ideas of voting theory originated more than 200 years ago, in the work of two French philosophers and mathematicians, Jean-Antoine the marquis de Condorcet and Jean-Charles de Borda. Both articulated a critique of plurality voting and proposed a (different) remedy. Their two methods are defined and illustrated in the examples below.

Plurality voting is, then as today, the most widely used voting method, of unrivaled simplicity. Each voter chooses one of the competing candidates, and the candidate with the largest support wins. Thus a voter only needs to designate his or her most preferred candidate.[1] Electors do not need to spell out a complete preference relation ordering all candidates, and the rule is at once transparent and easy to implement. German tribes elected their new chief by raising contenders on a shield, around which their supporters gathered: a simple head count and a couple of strong shields, is all the hardware they needed.

Condorcet and Borda agreed that plurality voting is seriously flawed, because it reflects only the distribution of the "top" candidates and fails to take into account the entire preference relation of the voters. It I vote for an extremist candidate who stands no chance of

1. Barring strategic manipulations, which cannot be ignored among "sophisticated" voters.

being elected, my ballot will play no role in the "real" contest between the two centrist candidates. Both the Condorcet and the Borda voting rule offer a (different) remedy to this difficulty.

The next three examples are borrowed from Borda's and Condorcet's essays on voting.

Example 4.1. Where Condorcet and Borda Agree Borda proposes the following example with 21 voters and three candidates (or outcomes) a, b, c. The voters are split in three groups of respective sizes 6, 7, and 8, and all voters within a given group have identical preferences:

Number of voters:	6	7	8
Top:	b	c	a
:	c	b	b
Bottom:	a	a	c

(2)

In the table above a preference relation is represented as a column, with the top candidate outcome ranked first and the bottom one ranked last. Note that a voter is never indifferent between two outcomes.

Plurality voting elects a in this profile, yet b is a more convincing compromise. Borda's and Condorcet's argument is that b is just below a in 8 ballots, but a is either just below or two positions below b in 13 ballots.

Borda proposes to tally the score of a candidate x by counting 2 points for each voter for whom x is the best candidate, 1 point for each voter who ranks x as second and 0 point for each voter ranking x last. Thus the Borda scores in example (2) are

$$\text{score }(a) = 8 \times 2 = 16$$

$$\text{score }(b) = 6 \times 2 + 15 \times 1 = 27$$

$$\text{score }(c) = 7 \times 2 + 6 \times 1 = 20$$

Now b has the highest Borda score hence is elected, whereas the plurality winner a has the lowest Borda score.

Condorcet's argument in support of the election of b is different. Suppose that the vote reduces to a duel between a and b (ignoring c altogether): then b wins by 13 votes against 8; similarly a duel between b and c has b winning 14 to 7, and finally c wins the a versus c duel by 13 to 8. Thus the *majority relation* R^m that records the winner of each duel x, y as "x is preferred to y by a majority of voters," is as follows:

$$bP^m c \quad bP^m a \quad cP^m a$$

We call b, the top outcome of the majority relation, the *Condorcet winner,* whereas a is the Condorcet *loser,* namely the bottom outcome of the majority relation.

In the second preference profile, 60 voters must decide among four candidates:

Number of voters:	23	19	18
Top:	a	b	c
	d	d	d
	b	a	a
Bottom:	c	c	b

The plurality winner is a but d is a more equitable compromise when the entire profile is taken into account.

Borda scores are computed by giving 3 points for each first place, 2 points for each second place, 1 point for each third place, and no point for each fourth place:

$$\text{score }(d) = 120 > \text{score }(a) = 106 > \text{score }(b) = 80 > \text{score }(c) = 54$$

As in example (2), the majority relation yields the same ordering of the candidates as the Borda scores:

$$dP^m a\ (37/23);\ \ dP^m b\ (41/19);\ \ dP^m c\ (42/18)$$

$$aP^m b\ (41/19);\ \ aP^m c\ (42/18)$$

$$bP^m c\ (42/18)$$

In our next example, the majority relation and the Borda scores make different recommendations, and this divergence reveals some important structural features of these two methods.

Example 4.2 Where Condorcet and Borda Disagree The profile has 26 voters and three candidates:

Number of voters:	15	11
	a	b
	b	c
	c	a

(3)

The plurality winner is a, and it is the Condorcet winner as well: a wins by 15 to 11 both duels against b and c. Borda's objection is that the eleven "minority" voters (supporters of b) dislike a more than the fifteen majority voters (supporters of a) dislike b. Indeed, a is the worst outcome for 42 percent of the voters, whereas b is always the first or second choice. This is reflected in their respective Borda scores:

$$\text{score }(b) = 15 + 2 \times 11 = 37 > \text{score }(a) = 2 \times 15 = 30 > \text{score }(c) = 11$$

Notice that Borda's argument relies on the conventional choice of points for first, second, and third place, which plays the role of a cardinal utility, albeit a mechanical one, unrelated to the real intensity of feelings of our voters for these three candidates (the voting rule prevents them from reporting any such intensity). Plurality voting is the voting method where first place gives one point and any other ranking gives no point.

The general family of *scoring methods* include Borda's and the plurality methods as special cases. Say that p candidates are competing. A scoring method is defined by the choice of a sequence of scores $s_1, s_2, \ldots, s_p$: a candidate x scores s_k points for each voter who ranks x in the kth place; the candidate or candidates with the highest total score wins. Naturally the scores decrease with respect to ranks, $s_1 \geq s_2 \geq \cdots \geq s_p$ and moreover $s_1 > s_p$, lest all candidates receive the same score irrespective of rank. Plurality corresponds to the scores $s_1 = 1, s_k = 0$ for $k = 2, \ldots, p$ and Borda to the scores $s_k = p - k$ for $k = 1, \ldots, p$. Another method of interest is *antiplurality,* for which $s_k = 1$ for $k = 1, \ldots, p-1$ and $s_p = 0$. In other words the antiplurality winner is the candidate who is least often regarded as the worst.

In example 4.2, depending on the choice of the scores, either a or b is elected[2]—but never c, whose score is always smaller than that of b, irrespective of the choice of s_1, s_2, s_3. This flexibility contrasts with the inflexible message of the Condorcet approach: a must be elected because the "will of the majority" is to prefer a to b and a to c. The fact that a is ranked last by the minority voters, whereas b is second best for all majority voters, is irrelevant to the majority relation. This is precisely the reason why the Borda method prefers b to a. Here Borda's method takes into account the entire preference profile but Condorcet's does not. We abstain at this point to make a normative judgment about this difference, which is at the heart of the axiom independence of irrelevant alternatives (IIA) discussed in section 4.6.

In our third example, slightly adapted from one of Condorcet's examples, the contrast between the Condorcet approach and the scoring approach (irrespective of the choice of scores) is especially clear.

Example 4.3 Condorcet against Scoring Methods There are 81 voters and three candidates, with the following preferences:

Number of voters:	30	3	25	14	9	
	a	a	b	b	c	
	b	c	a	c	a	(4)
	c	b	c	a	b	

2. For instance, $s_1 = 4, s_2 = 1, s_3 = 0$, makes a the winner with a score of 60 against 59 for b.

Candidate b is the plurality and the Borda winner in this profile. In fact, b wins for *any* choice of scores for first, second, and third place. To check this, we assume without loss of generality that these scores are 1, s, and 0, respectively, with $0 \leq s \leq 1$ (thus $s = 0$ is plurality and $s = \frac{1}{2}$ is the Borda method). Compute

$$\text{score }(b) = 39 + 30s > \text{score }(a) = 33 + 34s > \text{score }(c) = 9 + 17s$$

By contrast, the Condorcet winner is a because $aP^m b$ by 42/39 and $aP^m c$ by 58/23.

Notice that outcome c fares badly in both approaches. However, for the scoring methods the relative position of c with respect to b and a matters enormously: b wins essentially because c is much more often between b and a when b is first choice (this happens for 14 voters) than between a and b when a is first choice (happening only for 3 voters). On the other hand, a is a Condorcet winner, whether or not we take into account the irrelevant outcome c, or any other sure loser. This is, again, the IIA property alluded to above. It is a very strong argument in favor of the election of a in this example, and in support of the majority relation in general.

The most serious critique of the Condorcet approach is the observation, due to Condorcet himself, that the majority relation may cycle, meaning that it may fail the transitivity property (section 4.1). If the cycle involves the best outcomes of the majority relation, no Condorcet winner exists.

The simplest profiles of preferences exhibiting such cycles involve three outcomes a, b, c and only three different preference relations:

Number of voters:	n_1	n_2	n_3	
	a	c	b	
	b	a	c	(5)
	c	b	a	

If the sum of any two among the three numbers $n_i, i = 1, 2, 3,$ is greater than the third, the majority relation has the following cycle:

$$n_1 + n_2 > n_3 \Rightarrow aP^m b$$

$$n_1 + n_3 > n_2 \Rightarrow bP^m c$$

$$n_2 + n_3 > n_1 \Rightarrow cP^m a$$

and there is no Condorcet winner. Condorcet was keenly aware of this problem and proposed to break the cycle at his weakest link, namely to ignore the majority preference supported by the smallest majority.

For instance, suppose that $n_1 = 18, n_2 = 20$, and $n_3 = 10$. Then the link bP^mc is the weakest because b versus c yields a $28/20$ split versus $38/10$ for aP^mb and $30/18$ for cP^ma. Thus Condorcet suggests to elect c at this profile. Compare with the Borda method, electing a because score $(a) = 56 >$ score $(c) = 50 >$ score $(b) = 38$. In this example the election of either a or b is plausible.

Our last example uncovers a serious defect of any voting method electing the Condorcet winner whenever there is one, no matter how this method chooses to break the cycles of the majority relation when there is no Condorcet winner. To fix ideas, we assume as above that a cycle is broken at its weakest link, but the example can be adapted to any cycle-breaking rule.

Example 4.4 The Reunion Paradox We consider two disjoint groups of voters, with respectively 34 and 35 members, who vote over the same three candidates a, b, c. The first group contains left-handed voters, and the second one right-handed voters:

Number of left-handed voters:	10	6	6	12
	a	b	b	c
	b	a	c	a
	c	c	a	b

Number of right-handed voters:	18	17
	a	c
	c	a
	b	b

Candidate a is the majority winner among right-handed voters. Among left-handed voters, the majority relation has a cycle, $aP^mbP^mcP^ma$, of which the weakest link is cP^ma (by $18/16$ versus $22/12$ for the other two links); therefore we remove this link and elect a.

As a wins both among left-handed and among right-handed voters, we would expect—even request—that a be still declared the winner among the overall population of 69 voters. Yet c is the Condorcet winner there: cP^ma by $35/34$ and cP^mb by $47/22$!

The example above reveals a troubling paradoxical feature of the Condorcet approach. The paradox does not occur if a is a "real" Condorcet winner among left-handed and among right-handed electors, namely if for any other candidate x, a majority of lefties prefer a to x and a majority of righties do too: the union of these two majorities makes a majority in the grand population. Thus the paradox is a direct consequence of cycles in the majority relation.

Notice that any scoring method is immune to the reunion paradox. In the example a is the Borda winner in each subgroup and in the grand population. It is a simple exercise to check that for any system of scores, if two disjoint subsets of voters elect the same candidate a from the same pool of candidates, then a is still elected by the reunion of all the voters.

A related problem is the no-show paradox: at certain profiles of preferences, certain agents are better off staying home rather than participating in the election and casting a truthful ballot. This paradox affects all voting methods choosing the Condorcet winner when there is one, and none of the scoring methods.

4.3 Voting over Resource Allocation

In the discussion of the Condorcet and Borda voting methods, the set A of candidates/ outcomes is typically small, and voters may be endowed with arbitrary preferences over A. This is the correct modeling assumption when we speak about a political election, where the ability to report any ranking of the candidates is a basic individual right. However, when the issue on the ballot concerns the allocation of resources, some important restrictions on individual preferences come into play. The examples discussed below include voting time shares (example 4.5), over the location of a facility (example 4.6) and over tax- or surplus-sharing methods (examples 4.7 and 4.8). We find that majority voting works brilliantly in several of these problems (sections 4.4 and 4.5) but produces systematic cycling in others (example 4.5).

On the other hand, scoring methods are hopelessly impractical in all of these models because the set A of outcomes is large, and typically modeled as an infinite set, such as an interval of real numbers (example 4.6) or the simplex of an euclidean space (example 4.5). For instance, assume $A = [0, 1]$: a scoring method associates to an ordinal preference relation on A a scoring function representing the relation in question like a utility function. There are many different ways to define this representation,[3] and no natural way to select any of the scoring methods.

Another serious difficulty limiting the application of scoring methods in resource allocation problems comes from the IIA property: it is explained in the discussion following example 4.6.

In our next example, the issue is to divide a homogeneous private good when each voter cares only about his or her share. The commodity is a time share in example 4.5; it could be interpreted as money when the voters decide on a distribution of tax shares, or on the allocation of a surplus. The central feature is the pervasive cycling of the majority relation.

Example 4.5 Voting over Time Shares: Example 3.6a Continued We can choose any mixture $(x_1, \ldots, x_5)$ of the five radio stations, where x_i represents the time share of

3. One way is to pick a positive measure m on $[0, 1]$ and define the score $s(x) = m(P(x))$, where $P(x)$ is the set of outcomes y in A such that $x \succsim y$.

station i, and $\sum_1^5 x_i = 1$. The agents use majority voting to decide on the distribution of time shares.

The set of agents N is partitioned into five disjoint groups of one-minded fans: the agents in N_i like station i and no other station. We write n_i for the cardinality of N_i so that $\sum_1^5 n_i = n$. If one of the five subgroups N_i contains a majority of voters—$n_i > n/2$—then playing their station all the time—meaning $x_i = 1$—is the Condorcet winner (and the plurality winner as well). On the other hand, if none of the five coalitions forms an absolute majority, $n_i < n/2$ for $i = 1, \ldots, 5$, then the majority relation is strongly cyclic and there is no Condorcet winner.

Consider an arbitrary distribution of time shares $x_1, \ldots, x_5$. Suppose that station 1 receives a positive share $x_1 > 0$. The coalition of all agents who do not like station 1 can "gang up" on the n_1 supporters of this station and give a positive piece of the spoil to each one of the four other stations. In other words, consider the following vector of timeshares $y_1, \ldots, y_5$:

$$y_1 = 0, \quad y_i = x_i + \tfrac{1}{4}x_1, \qquad \text{for } i = 2, 3, 4, 5$$

Every supporter of station $i, i = 2, \ldots, 5$, strictly prefers the distribution y over x. Our assumption $n_1 < n/2$ means that $\sum_2^5 n_i > n/2$; that is to say, a majority of voters prefer y over x so that x can't be a Condorcet winner. But, for *any* distribution x of time shares, some station i receives a positive share, and the argument above shows that taking away the share x_i to distribute it among all other stations is a move from which a majority of voters benefit. Hence there is no Condorcet winner.

The example illustrates a strategic situation known as "destructive competition," that often emerges when relatively small coalitions can inflict severe negative externalities upon the complementary coalition. Examples of destructive competition involving production and exchange of private goods are discussed in section 7.3. There as here, the issue is a failure of the logic of private contracting. Every distribution of time shares among the five coalitions is threatened by a private contract of at most four coalitions joining to deprive the remaining coalition of any benefit whatsoever. The cycles of the majority relation correspond to the never ending process of these majority "coups." Instability and unpredictability of the eventual outcome is a consequence of the excessive power awarded to any majority of the voters. A solution to destructive competition in the voting context is to reduce the power of coalitions, for instance, by requiring a qualified majority (a larger support) to overturn a given outcome.

In example 4.5, to fix ideas, assume that $n = 100$ and $n_1 = 40, n_2 = 25, n_3 = 15, n_4 = 12, n_5 = 8$. We require a qualified majority of Q or more to overturn any given allocation. If $51 \leq Q \leq 60$, destructive competition reigns, exactly as before, because any reunion of four out of the five homogeneous subgroups reaches the quota Q. If $61 \leq Q \leq 75$, the coalition $N_2 \cup N_3 \cup N_4 \cup N_5$ no longer passes the quota Q, so these agents can't get together to "steal" the time share x_1. On the other hand, the four other coalitions made of four of the five subgroups reaches Q; therefore the argument in example 4.5 shows that a distribution x where $x_i > 0$ for one of $i = 2, 3, 4, 5$ will be outvoted when the union of N_j, for all $j \neq i$, forms. In turn this establishes that the only stable allocation of time shares is $x_1 = 1, x_i = 0$ for $i = 2, 3, 4, 5$! The homogeneous group N_1 holds veto power and uses it to extract the entire surplus. The strategic logic here is core stability (as in sections 7.1 through 7.3).

Next consider the case $76 \leq Q \leq 85$. Now we need some voters in N_1 and some voters in N_2 to form a coalition of size Q or more; therefore both N_1 and N_2 have veto power. As a result the corestable outcomes are all distributions x of time shares such that $x_1 + x_2 = 1$, $x_i = 0$ for $i = 3, 4, 5$. A similar argument shows that all distributions x such that $x_1 + x_2 + x_3 = 1$, and only those, are corestable when $86 \leq Q \leq 88$. Finally when the quota reaches 93, the core stability property loses all bite, and any distribution x, $\sum_1^5 x_i = 1$, is stable.

4.4 Single-Peaked Preferences

The domain of preferences discussed in this section guarantees the transitivity of the majority relation, in turn making the Condorcet approach to voting unambiguously successful.

Example 4.6 Location of a Facility (Example 3.4 Continued) As in example 3.4 the voters live in a linear city represented as the interval [0, 1]. A voter living at x, $0 \leq x \leq 1$, wishes that the facility be located as close as possible to x, and her utility when the facility is at y is the negative of the distance between x and y, $u_i(y) = -|y - x_i|$. The distribution of our voters along [0, 1] is represented by a cumulative function F, where $F(z)$ is the proportion of voters living on $[0, x]$, and $1 - F(z)$ is the proportion of those living on $[z, 1]$.

We assume, for simplicity, that there is a large number of voters spread continuously between 0 and 1 so that the function F increases continuously from $F(0) = 0$ to $F(1) = 1$. In other words, the proportion of agents living at a given point z is always zero.[4]

The median of the distribution F is this point y^* such that $F(y^*) = \frac{1}{2}$, meaning that half of the population lives to the left of y^* and half to its right. Recall from example 3.4 that y^* is the classical utilitarian solution. In fact, y^* is the Condorcet winner as well.

If we compare y^* to y on its left, $0 \leq y < y^*$, all voters in $[y^*, 1]$ prefer y^* to y, and so do those in $[(y + y^*)/2, y^*]$ because they are closer to y^* than to y. Thus the supporters of

4. All results are preserved if we deal with a small finite set of voters or if a positive fraction of the voters are piled up at certain locations.

y^* versus y form the proportion $1 - F((y + y^*)/2)$ of the population, and this constitutes a majority:

$$y < y^* \Rightarrow F\left(\frac{y + y^*}{2}\right) < F(y^*) = \frac{1}{2} \Rightarrow 1 - F\left(\frac{y + y^*}{2}\right) > \frac{1}{2}$$

A symmetrical argument applies when we compare y^* to y in $]y^*, 1]$: all voters living in $[0, (y + y^*)/2]$ prefer y^* to y and they form a majority because $F((y + y^*)/2) > F(y^*) = \frac{1}{2}$. This proves our claim that y^* is the majority winner. By similar arguments it is easy to show that the majority relation coincides with the preferences of an agent living at y^*. Location y is preferred by a majority to location y' if and only if y is closer to y', namely $|y - y^*| < |y' - y^*|$ (see exercise 4.6).

The remarkable coincidence of the Condorcet winner and the utilitarian optimum in example 4.6 depends on the particular assumption that the distance from the facility to one's home is the disutility function of each agent. An important observation is that the median of the distribution is still the Condorcet winner (if not the utilitarian optimum) for a much larger domain of individual preferences, called the *single-peaked* preferences.

Given an ordering of the set A from left to right (from 0 to 1 if outcomes in A are represented by real numbers as in the example above), we write $x < y$ when x is to the left of y and we say that z is "between" x and y if either $x \le z \le y$ or $y \le z \le x$. The preference relation R_i is *single peaked* (in the ordering of A) *with peak* x^i if x^i is the top outcome of R_i in A, and moreover for all outcome x, $x \neq x^i$, R_i prefers any outcome between x^i and x to x itself.

The simple geometric intuition for single-peaked preferences is shown on figure 4.1, where A is represented by an interval $[a, b]$. The preferences are increasing when x increases (moves right) from a to the peak x^i of R_i; they are decreasing when x increases from x^i to b. The important point is that the comparison of outcomes across the peak—say x to its left versus y to its right—are not restricted: see figure 4.1.

The assumption that all individual preferences are single-peaked is plausible in many problems where the outcomes are naturally arranged along a line. This is especially clear if we vote over the drinking age, or the tax rate, or the length of a patent. Another important example is the Downsian model of political competition, where A models the size of the defense budget, the funding of public education, and so on. Of course, a real assembly is rarely so simple as to be lined up from leftist to rightist on all issues, but on specific issues the assumption makes sense. A final example is product differentiation: the group of agents N must pick a software or a copier, or any item whose cost is equally split among all. There is a single dimension of "quality," ordered by price. The assumption that each user

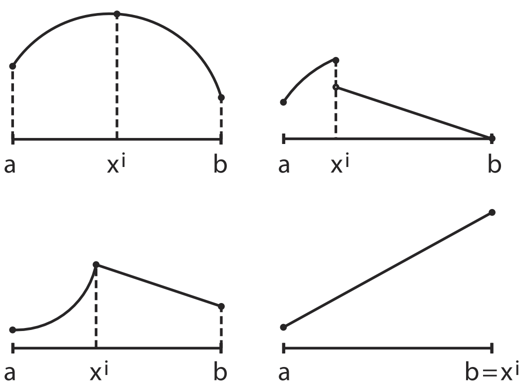

Figure 4.1
Single-peaked preferences

has single-peaked preferences over the different levels of quality amounts to the familiar convexity property.[5]

In example 4.6 the central argument is that the median y^* of the distribution of individual peaks x^i (where agent i lives) is the Condorcet winner. This argument is valid for any profile of single-peaked preferences. Consider an outcome x to the left of y^* $(x < y^*)$: all agents whose peak x^i is to the right of y^* or at y^* prefer y^* to x because y^* is between x and x^i. These agents form a majority by definition of the median; therefore y^* defeats x for the majority relation. A symmetrical argument shows that y^* defeats all outcomes to its right.

Another property of example 4.6 is preserved under any profile of single-peaked preferences: the majority relation is transitive, and is single peaked as well. Its peak is the median peak, and so is the Condorcet winner. Exercise 4.6 explains these facts.

The Condorcet winner is particularly easy to implement when preferences are single-peaked because each agent only needs to report her peak. The way she compares outcomes on the left of her peak to outcomes on its right, does not affect the Condorcet winner (although it does affect the majority relation over outcomes below the winner).

Moreover the definition of the feasible set A far away from the Condorcet winner does not matter either. If we extend A to the right or to the left by adding a few outcomes that stand no chance of being elected, the Condorcet winner does not change, another illustration of the property Independence of Irrelevant alternatives. Here is a simple example with seven agents with single-peaked preferences on [0, 100] and the following peaks:

$$x^1 = 35 \quad x^2 = 10 \quad x^3 = 22 \quad x^4 = 78 \quad x^5 = 92 \quad x^6 = 18 \quad x^7 = 50$$

5. If level x is preferred to level y, and x' is also preferred to y, then all levels in $[x, x']$ are preferred to y.

On $A = [0, 100]$ the median of these peaks is at $y^* = 35$, and this is true on the following smaller intervals:

$$B = [20, 75]: \quad \text{peaks } x^1 = 35,\ \tilde{x}^2 = \tilde{x}^6 = 20,\ x^3 = 22,\ \tilde{x}^4 = \tilde{x}^5 = 75,\ x^7 = 50$$

$$C = [20, 40]: \quad \text{peaks } x^1 = 35,\ \tilde{x}^2 = \tilde{x}^6 = 20,\ x^3 = 22,\ \tilde{x}^4 = \tilde{x}^5 = \tilde{x}^7 = 40$$

and on any interval containing 35.

By contrast, if one wishes to apply a scoring method in a problem like example 4.6, the entire preference relation of every voter matters, and so do the precise end points of the set A.

The last but not least desirable feature of the Condorcet method is *strategy-proofness:* a voter has no incentive to lie strategically when reporting the peak of his preferences. Even when a group of voters attempt to jointly misreport their peaks, they cannot find a move from which they all benefit.

We check this claim in example 4.6, where y^* is the median of the distribution of individual peaks x^i. In the argument below, it does not matter whether the set of agents is large (infinite, as in example 4.6) or small, finite, as in the numerical example three paragraphs above.

Denote by N_- the set of agents whose peak is (strictly) to the left of y^* on A $(x^i < y^*)$, by N_+ the set of those whose peak is to its right $(y^* < x^i)$, and by N_0 those with $x^i = y^*$. Suppose that the coalition T of voters agree to alter their reported peaks, from the true peak x^i to a fake $\tilde{x}^i$, while the rest of the agents report their peak truthfully as before. Denote by z^* the new median of the reported peaks: we show that either $z^* = y^*$ or at least one agent in T strictly prefers y^* to z^*. In the former case, the joint misreport is inconsequential; in the latter, it is not plausible because participation has to be voluntary.

The proof is by contradiction. Suppose that $z^* \neq y^*$ and that no i in T strictly prefers y^* to z^*. Say that z^* is to the right of y^* in A $(y^* < z^*)$. Because preferences are single peaked, everyone in N_- and in N_0 strictly prefers y^* to z^*; therefore T is contained in N_+. By definition of the median, $N_- \cup N_0$ forms a strict majority, and we just proved that they all still report their true peak; therefore a majority prefers y^* to z^*, and z^* cannot be chosen when T misreports, contradiction. A symmetrical argument applies when z^* is to the left of y^*.

Strategy-proofness is the ultimate test of incentive-compatibility in mechanism design. In a strategy-proof allocation or voting mechanism, no participant has any incentive to report his own characteristics (preferences, endowment) strategically: the simple truth is always my best move, whether I have no information about other agents' messages, or full information, or anything in between. Two very important examples of strategy-proof mechanisms are majority voting over single-peaked preferences, and the competitive equilibrium when each market participant is negligible with respect to the total endowment of the economy (sections 6.3 and 7.1).

Note that all scoring rules fail to be strategy-proof, even when individual preferences are single peaked. An example with three outcomes a, b, c and nine voters illustrate the point. We have five voters with preferences $a \succ b \succ c$, and four voters with $b \succ a \succ c$; these preferences are single peaked for the ordering $a < b < c$. The Borda winner is a, but if the four "losing" voters report $b \succ c \succ a$, the winner is b and the misreport is thus profitable.

Majority voting à la Condorcet is thus a compelling voting method when the outcomes are arranged along a line and individual preferences are single-peaked. The assumption that the outcome set is a one-dimensional interval can be weakened to a tree pattern, but not to a one-dimensional "loop"; see exercises 4.8 and 4.9.

Real life voting rules, however, do not impose any restriction on the shape of individual preferences, and in this case incentive-compatibility becomes a thorny issue. A disappointing impossibility result, discovered in the early 1970s, eliminates any hope of a simple answer.

Any voting method defined for *all* rational preferences over a set A of three or more outcomes must *fail* the strategyproofness property: at some preference profile, some agent will be able to "rig" the election to his or her advantage (i.e., bring about the election of a better outcome) by reporting untruthfully his or her preferences. This important fact, known as the Gibbard-Satterthwaite theorem, is technically equivalent to Arrow's theorem discussed in section 4.6. It is formally stated in chapter 8.

4.5 Intermediate Preferences

We turn to the second configuration of preferences guaranteeing that the majority relation is transitive, hence a Condorcet winner exists. The property of *intermediate preferences* relies on an ordering of the agents, instead of an ordering of the outcomes in the case of single-peaked preferences. In example 4.8 below, the agents choose a taxation method and differ only by their pre-tax income (they are selfish, only interested in maximizing aftertax income): they are naturally ordered along the income scale.

We say that the profile has the intermediate preferences property if whenever two agents i, j agree to prefer outcome a to b, so do all agents "between" i and j. Say that the 100 agents are ordered as $N = \{1, 2, \ldots, 100\}$. Intermediate preferences imply that the set $N(a, b)$ of agents preferring a over b is an interval $[i_1, i_2]$, namely $N(a, b)$ consists of all agents i such that $i_1 \leq i \leq i_2$. The same observation applies to the set $N(b, a)$ of agents preferring b to a. Barring indifferences for simplicity, we see that $N(a, b)$ and $N(b, a)$ partition $[1, 100]$ in two disjoint intervals: thus $N(a, b)$ must be an interval of the type $[1,\ i^*]$ or $[j^*, 100]$.

We check that the majority relation is transitive. Pick three outcomes a, b, c such that $N(a, b)$ and $N(b, c)$ each contain 51 agents or more, so that the majority relation prefers a to b and b to c. If $N(a, b) = [1, i^*]$ and $N(b, c) = [1, j^*]$, with $i^*, j^* \geq 51$, then all agents in

$[1, 51]$ prefer a to b and b to c, hence a to c, and the majority relation prefers a to c as claimed. A symmetrical argument applies if $N(a, b) = [i^*, 100]$ and $N(b, c) = [j^*, 100]$, with $i^*, j^* \leq 50$. Suppose next that $N(a, b) = [1, i^*]$ and $N(b, c) = [j^*, 100]$, with $i^* \geq 51$ and $j^* \leq 50$. Then agents 50 and 51 both belong to $N(a, b)$ and $N(b, c)$, hence to $N(a, c)$ as well. If $N(a, c)$ takes the form $[1, i]$, this implies $i \geq 51$ and if $N(a, c) = [j, 100]$, this implies $j \leq 50$: in both cases $N(a, c)$ is a strict majority and the claim is proved.

Example 4.7 Voting over Three Surplus-Sharing Methods We consider first the three surplus-sharing methods in section 2.2, namely the proportional (PRO), equal surplus (ES), and uniform gains (UG) methods.

Given a particular profile of claims and amount of resources t to be divided, our agents vote to choose which method will be implemented. Agents are ranked by the size of their initial claims/investments. They compare the three methods exclusively by the size of their own share of total surplus, and the larger the better. Here is an example with 11 voters, total resources $t = 745$, and initial claims ranging from 10 to 120 and totalling 580. We compute the shares allocated by our three methods:

Agent	1	2	3	4	5	6	7	8	9	10	11
Claim	10	10	20	25	40	40	60	70	85	100	120
PRO	12.8	12.8	25.7	32.11	51.4	51.4	77.1	89.9	109.2	128.4	154.1
ES	25	25	35	40	55	55	75	85	100	115	135
UG	51.7	51.7	51.7	51.7	51.7	51.7	60	70	85	100	120

The four agents with the smallest claims rank the uniform method above equal surplus and the latter above proportional. The five agents with the largest claims have the exactly opposite preferences.[6] For the two middle agents with claim 40, the best method is equal surplus. Thus all preferences are single peaked with respect to the ordering {uniform, equal surplus, proportional} of the three outcomes. The median peak is 40 and "equal surplus" is the Condorcet winner at this profile.

Next we check the intermediate preferences property. It suffices to check that the sets $N(a, b)$ are intervals of the form $[1, i]$ or $[j, 11]$ for all a, b:

$$N(\text{UG}, \text{PRO}) = N(\text{ES}, \text{PRO}) = [1, 6]; \quad N(\text{ES}, \text{UG}) = [5, 11]$$

Assume that the commodity being distributed is a "bad," and that individual claims represent a liability. Now an agent prefers method m to another method m' if and only if his share under m is smaller than under m'. Observe that the new preferences are not single peaked

6. Recall from exercise 2.5 that the smallest claim and largest claim agents have these preferences for *all* surplus-sharing problems.

anymore, because each method is the worst outcome for some agents: PRO for agents $[7, 11]$, ES for 5, 6, and UG for $[1, 4]$. On the other hand, the intermediate preferences property is preserved because the intervals $N(a, b)$ and $N(b, a)$ are simply exchanged:

$$N'(\text{UG, PRO}) = N'(\text{ES, PRO}) = [7, 11]; \quad N'(\text{ES, UG}) = [1, 4]$$

The proportional solution is the Condorcet winner; indeed, it is the best method for the six agents $1, 2, \ldots, 6$.

Exercise 4.13 shows that for all surplus-sharing problems of any size, the preferences over the three methods PRO, ES and UG have the intermediate preference property. Our next example shows that the same holds true when agents choose a rationing method within the one-parameter family uncovered in section 2.4.

Example 4.8 Voting over Tax Schedules The agents choose one of the equal sacrifice methods described at the end of section 2.4, namely they choose a common utility function u to measure sacrifice, and this function takes one of the following forms:

$$\begin{aligned} u_p(z) &= -\frac{1}{z^p} \quad \text{for some positive parameter } p \\ u_0(z) &= v_0(z) = \log\, z \\ v_q(z) &= z^q \quad \text{for some positive parameter } q \end{aligned} \tag{6}$$

Once their vote has elected one such utility function, taxes are computed by solving system (10) in chapter 2. Recall that the function $u_0 = v_0$ corresponds to the proportional method (flat tax), the function v_1 to uniform losses (head tax), and u_∞ to uniform gains (full redistribution).

Remarkably, the intermediate preferences property holds true for any number of taxpayers and any profile of taxable incomes, so that majority voting always delivers a Condorcet winner. Before explaining this result, we note that the pattern of preferences over this family of tax schedules is not in general single peaked for any ordering of the family.

Consider the profile of taxable incomes $x_1 = 20$, $x_2 = 80$, $x_3 = 100$, and total aftertax income $t = 120$ (i.e., total tax levied is 80). The three basic methods give the following aftertax incomes:

Agent	1	2	3
PRO	12	48	60
UG	20	50	50
UL	0	50	70

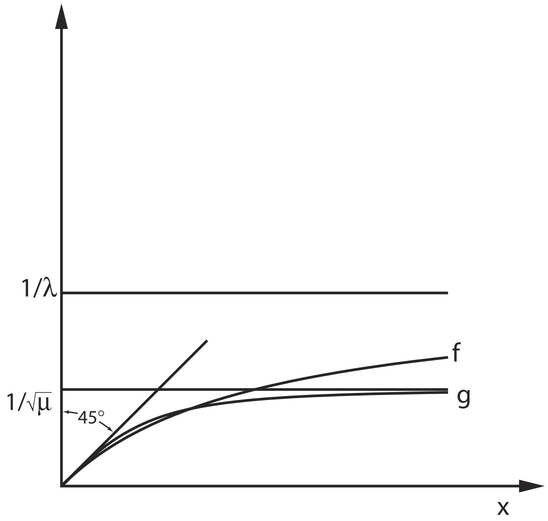

Figure 4.2a
Single crossing property in example 4.8

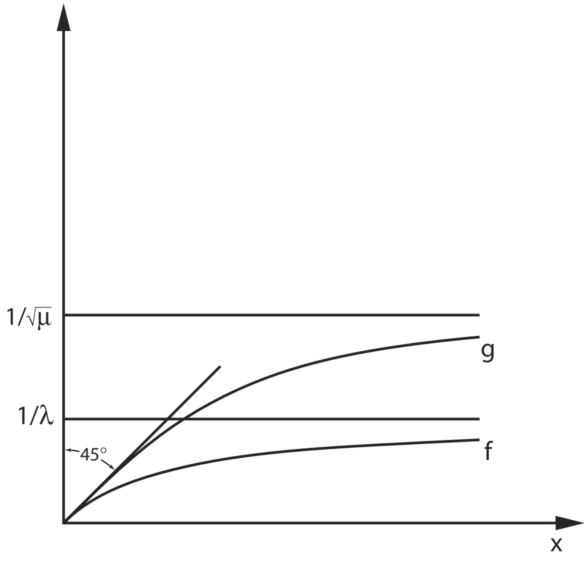

Figure 4.2b
Single crossing property in example 4.8

so that PRO is the worst method for agent 2, UG is the worst for agent 3, and UL is the worst for agent 1, which rules out single-peakedness for any ordering of the three methods.

We check now the intermediate preference property for the two methods u_1 and u_2.[7] Recall that the u_p methods are computed by solving the simpler system (9) in chapter 2. Fix the profile of taxable incomes x_i, the total aftertax income t, and denote by y_i, z_i the aftertax incomes under u_1 and u_2 respectively. System (9) tells us that there are two positive numbers λ, μ such that

$$\frac{1}{y_i} - \frac{1}{x_i} = \lambda \quad \text{and} \quad \frac{1}{z_i^2} - \frac{1}{x_i^2} = \mu \quad \text{for all } i$$

$$\Leftrightarrow y_i = \frac{x_i}{1 + \lambda x_i} \quad \text{and} \quad z_i = \frac{x_i}{\left(1 + \mu x_i^2\right)^{1/2}} \quad \text{for all } i$$

Compare the two concave functions $f(x) = x/(1 + \lambda x)$ and $\delta(x) = x/(1 + \mu x^2)^{1/2}$ for $x \geq 0$: they coincide at $x = 0$, where they both have a slope of 1, then they cross once at $x = 2\lambda/(\mu - \lambda^2)$ if $\mu > \lambda^2$, or not at all if $\mu \leq \lambda^2$. Figure 4.2 illustrate these two cases.

Therefore for any λ and μ, the set of agents i for whom $f(x_i) > g(x_i)$ is made either of all agents whose taxable income is below a certain level, or of all those with income above some level. This is exactly the intermediate preferences property when we order the agents according to their taxable incomes.

4.6 Preference Aggregation and Arrow's Theorem

As indicated in section 4.1, a social choice problem is made of three ingredients. The set A contains all the outcomes (states of the world) among which the set N of concerned agents—the "society"—must choose one. The choice of an outcome a in A affects the ordinal welfare of each concerned agent i: this is captured by a preference relation R_i on A, namely a complete and transitive binary relation.

The differences between individual preferences are resolved by the *aggregation* method F, associating to each profile of preferences $\tilde{R} = (R_i, i \in N)$, a collective—or social—preference relation $R^* = F(\tilde{R})$, interpreted as the ordinal collective welfare. The aggregation method plays exactly the same role, in the ordinal context, as the collective utility function in the cardinal context of the previous chapter.

7. The general argument is the subject of exercise 4.16.

Collective welfare is identified with a preference relation R^*, guiding the collective choice over any subset B of feasible outcomes: welfare is identical to choice. This construction is anthropomorphic, in the sense that the collective body is treated exactly like an individual agent. The mechanical computation of the collective relation R^* from the profile $\tilde{R}$ is social engineering at its best, or its worst, namely a controversial normative construction. It suggests one way of thinking about democratic institutions, but competing models offer alternative answers to this central question of political philosophy.

Recall from section 1.5 the basic tenet of the minimal state (libertarian) doctrine: collective decisions merely result from the interaction of free citizens exercising their political rights. This decision process may indeed yield a pattern of choices that can in no way be deciphered as rational, as maximizing some underlying collective preference. But this feature is by no means a subject of concern: collective choice is devoid of normative content, and any outcome of the free interplay of individual rights is as good as any other.

Social choice theory takes the diametrically opposed view that the process to reach a democratic compromise should rest on sound axiomatic foundations and allow positive predictions. For instance, cycles of the majority relation are deemed undesirable because they lead to the chronic formation of unstable coalitions and arbitrariness of the final decision which, ultimately, threatens the political legitimacy of the institutions for collective decision. The model of preference aggregation is the most general—the most ambitious—project of mechanism design in the microeconomic tradition. Its limited success, underlined by Arrow's impossibility theorem, can just as easily be viewed as a vindication of the minimal state doctrine—the search for rationality of collective choice is hopeless—or as the first step in a larger project of social engineering poised to discover specific allocation problems for which rational collective choice is within our reach.

The two voting methods proposed by Condorcet and Borda suggest two simple aggregation methods. Condorcet's argument is that for a given profile $\tilde{R}$ the majority relation is the correct expression of the general will (*volonté générale*). Formally this relation $R^m = F(\tilde{R})$ is defined as follows, for any pair of outcomes $x,\ y$:

$$x R^m y \Leftrightarrow |\{i \in N \mid x P_i y\}| \geq |\{i \in N \mid y P_i x\}| \tag{7}$$

namely the supporters of x against y are not outnumbered by their opponents.

We saw in the previous section that for some preference profiles $\tilde{R}$, the relation R^m is cyclic, hence violates the transitivity requirement for rational choice. In many collective decision problems, we can exclude a priori no preference relation on A. Voting over candidates to a political office is an obvious example, because of freedom of opinions. In this case the majority relation is not a valid aggregation method. On the other hand, the Borda scoring method provides an aggregation method for all preference profiles on any finite set A.

If p is the number of outcomes in A and R_i an arbitrary preference relation on A, we define the Borda scores $s(a, R_i)$ awarded by R_i. If R_i expresses strict preferences between any two outcomes (agent i is never indifferent between a and b), we set, as in section 4.2,

$$s(a, R_i) = p - k \quad \text{for the outcome ranked} \quad k\text{th in } A$$

(hence the top outcome gets $p - 1$ points and the bottom one gets 0 point). When R_i is indifferent between say a, b and c, these outcomes split equally the total scores they would fetch if preferences were strict and a, b, c were adjacent. To illustrate this straightforward construction, consider $p = 8$ and the following preference R_i:

$$\{a, b\} P_i c P_i \{d, e, f\} P_i g P_i h$$

where the brackets denote an indifference class; for example, R_i is indifferent toward d, e, and f. Outcomes a and b split the total score $7 + 6$, so they get 6.5 each; c gets 5; d, e, and f each get 3; g gets 1, and h gets 0.

The Borda aggregation method yields $R^b = F(\tilde{R})$ as follows for any pair of outcomes x, y:

$$x R^b y \Leftrightarrow \sum_{i \in N} s(x, R_i) \geq \sum_{i \in N} s(y, R_i) \tag{8}$$

The transitivity of the relation R^b follows at once from that of the inequality relation between scores.

Consider the profile (4) in example 4.3. We noted there that b, the Borda winner, owes its success over a to the relative position of c vis-à-vis b and a. By contrast, the majority relation (transitive in this example) puts a above b: if the issue is to choose between a and b, it compares the numbers of supporters of a versus b and b versus a, without paying any attention to c at all.

The property *independence of irrelevant alternatives* (IIA) requires that the collective preference R^* between any two outcomes x and y only depend upon individual preferences between any two outcomes. That is, if $\tilde{R}$ and $\tilde{R}'$ are two profiles of preferences that produce exactly the same sets of supporters of x versus y and y versus x:

$$\text{for all } i: \quad x P_i y \Leftrightarrow x P_i' y, \quad y P_i x \Leftrightarrow y P_i' x, \quad x I_i y \Leftrightarrow x I_i' y$$

then the collective preferences $R = F(\tilde{R})$ and $R' = F(\tilde{R}')$ compare x and y in precisely the same way:

$$x P y \Leftrightarrow x P' y, \quad y P x \Leftrightarrow y P' x, \quad x I y \Leftrightarrow x I' y$$

The majority aggregation method (7) meets the IIA property but does not always produce a rational collective preference. When it cycles, say $a P^m b$, $b P^m c$, and $c P^m a$, it is helpless to guide the choice among a, b, and c. Any method to break the cycle, for instance, at its weakest link (see example 4.4 and the discussion preceding it), leads to a violation of IIA: if we declare a the winner among a, b, c because the statement $c P^m a$ is the "weakest" (has the smallest number of supporters), then "a wins over b" depends on the individual preferences over a, b, and c, not just over a and b. Similarly, if the collective preference declares a, b, and c to be indifferent, then the mere knowledge of the supporters of a versus b and b versus a is not enough to determine the collective preference between a and b: it also matters whether a, b are part of some cycle of the majority relation (involving other outcomes) or not.

The Borda aggregation method (8) produces a rational collective preference for any profile $\tilde{R}$ but fails the IIA property. This means that the choice of the set A of outcomes/candidates on the ballot is of critical importance to the eventual outcome. Adding to $A = \{a, b\}$ a candidate c who stands no chance to win against either a or b (and will be ranked last by the collective preference) may turn around the choice between a and b, as in examples 4.2 and 4.3. Control of the "agenda," namely the set A of eligible candidates, is often tantamount to control of the election and of the collective preference. Thus the definition of A is controversial and may be the subject of a preliminary round of collective decision-making, the agenda of which is itself a matter of dispute, and so on ad infinitum.

Arrow's impossibility theorem explores the sharp trade-offs between the IAA property and the rationality of collective preferences, in the formal context of aggregation functions. In a nutshell, the theorem says that *any* aggregation function producing a rational collective preference and meeting IIA must be highly undesirable on account of its lack of efficiency or of fairness. A formal statement is given in chapter 8.

For instance, suppose that we want an efficient aggregation method, namely we insist that the collective preference R^* respects the unanimous preferences of the citizens. For any two outcomes x, y,

$$\{x P_i y \text{ for all } i \in N\} \Rightarrow x P^* y$$

Then the only rational aggregation methods meeting IIA are the *dictatorial* methods, where the collective preference relation R^* coincides with R_i^*, the preference relation of the dictator i^*. The point is that the dictator's preferences prevail irrespective of those of the rest of the agents, a state of affairs that we may call maximally unfair.

Suppose next that we restrict our attention to aggregation functions that are fair in the sense that all voters have equal influence a priori on the collective preference.[8] Some unpalatable rational aggregation methods meeting IIA are the "imposed" methods, always selecting the same collective preference R_o, irrespective of the preferences of our citizens. Such a method is fair, but it is pathetically inefficient: even when the citizens share a common

8. Formally, this means that the function F is symmetrical in its n variables R_i.

preference $R_i = R^o$ for all i, the collective preference ignores this fact entirely. It turns out that the only fair rational aggregation methods meeting IIA are imposed, except for at most two fixed outcomes a, b that can be compared, for instance, by the majority relation.

The proof of Arrow's theorem is beyond the scope of this book. The two "ways out" of the impossibility have been discussed earlier in this chapter. One way is to restrict the domain of admissible preference profiles, as with the single peaked (section 4.3) or intermediate (section 4.4) preferences. Another way is to weaken the rationality properties of the collective preference, by only requiring that its strict preferences do not cycle. This is the idea of qualified majorities briefly discussed at the end of section 4.3. The approach leads to indecisive collective preferences, however, with too many different outcomes declared winners.

4.7 Introduction to the Literature

The theory of rational choice and ordinal preferences, discussed in section 4.1, is the very foundation of microeconomic analysis. Textbook presentations are easy to find; a particularly good one is in Mas-Colell, Whinston, and Green (1995, ch. 1).

The original discussion of the optimal design of voting rules by Condorcet (dating back to 1785), is still extremely useful and accessible in the excellent English translation of McLean and Urken (1995). Our examples 4.1 to 4.3 are adapted from Condorcet's original examples.

The reunion paradox (example 4.4) is the basis of an important result due to Smith (1973) and Young (1974): the scoring methods are the only voting methods avoiding the reunion paradox and treating symmetrically both voters and candidates. Two excellent surveys on voting rules are Brams (1994) and Brams and Fishburn (2002); see also Ordershook (1986).

The central result underlying section 4.4 is due to Black (1948) and is formally described in exercise 4.6: the majority relation is a rational preference when individual preferences are single-peaked. The related property of strategyproofness has been studied extensively: Moulin (1980) uncovers the full family of strategy-proof voting rules on the single-peaked domain, a result later extended to multidimensional versions of this domain; see Sprumont (1995) and Barbera (2001) for a survey of this literature.

The seminal impossibility result on strategy-proof voting with unrestricted preferences is due to Gibbard (1973) and Satterthwaite (1975): it is briefly discussed at the end of section 4.4. Textbook presentations can be found in Moulin (1988, ch. 10) and Mas-Colell, Whinston, and Green (1995).

Grandmont (1978) introduced the notion of intermediate preferences, section 4.5, and Roberts (1977) noticed that preferences over tax schedules can be expected to fit this pattern.

Arrow's 1951 book, with the original proof of his theorem, is the unambiguous starting point of mathematical welfare economics, and in this sense its influence pervades the entire book. The formal statement and proof of the theorem is available in many books, among them Sen (1970), Kelly (1978), Moulin (1988), and Mas-Colell, Whinston, and Green (1995).

Peremans and Storcken (1996) introduce the concept of single-dipped preferences discussed in exercise 4.7.

Exercises to Chapter 4

Exercise 4.1 Two More Examples of Condorcet

a. In the following profile with 60 voters shows that the majority relation and the Borda scores yield the same ranking of the three outcomes. Compare it with the ranking of plurality voting.

Number of voters:	18	5	16	3	13	5
	a	*a*	*b*	*b*	*c*	*c*
	c	*b*	*c*	*a*	*b*	*a*
	b	*c*	*a*	*c*	*a*	*b*

b. In the next profile the majority relation has a cycle. What outcome wins if we use Condorcet's idea to break the cycle at its weakest link? What outcome (outcomes) is (are) elected by *some* scoring method?

Number of voters:	23	2	17	10	8
	a	*b*	*b*	*c*	*c*
	b	*a*	*c*	*a*	*b*
	c	*c*	*a*	*b*	*a*

Exercise 4.2 An Example due to Joe Malkovitch

We have 55 voters and five outcomes. The profile of preferences is as follows:

Number of voters:	18	12	10	9	4	2
	a	*b*	*c*	*d*	*e*	*e*
	e	*e*	*b*	*c*	*b*	*c*
	d	*d*	*e*	*e*	*d*	*d*
	c	*c*	*d*	*b*	*c*	*b*
	b	*a*	*a*	*a*	*a*	*a*

Check that the majority relation is transitive and orders the candidates exactly like the Borda scores. Compare with plurality voting.

Exercise 4.3 More on Profile (5)

Consider the preference profile (5) where the majority relation has a cycle: $n_i < n/2$ for $i = 1, 2, 3$. Check that the Borda winner corresponds to the index i maximizing $n_i - n_{i+1}$, where we set $n_{3+1} = n_1$. Check that the Condorcet solution (break the cycle at its weakest link) corresponds to i maximizing n_i.

Give an example where these two solutions are unique and different.

Exercise 4.4

Consider the profile with seven voters and four candidates:

Number of voters:	3	2	2
	c	b	a
	b	a	d
	a	d	c
	d	c	b

a. Compute the majority relation and show that it has several cycles. By "breaking" the two weakest majority preferences, check that the ordering $cPbPaPd$ obtains.

b. Compute the Borda ranking of our four candidates. Find an outcome x such that, upon removing x, the ranking of the other three candidates is completely reversed! The violation of which axiom does this illustrate?

Exercise 4.5 Location of a Noxious Facility (Example 3.5)

The public facility that must be located somewhere in [0, 1] is undesirable (prison, waste disposal) and the distance from agent i's home x^i to the site y of the facility measures her utility $u_i(y) = |x^i - y|$. See example 3.5.

a. Show that the endpoint (0 or 1) farthest away from the median y^* of the distribution F of agents (see example 4.6) is the Condorcet winner. Recall from example 3.5 that the utilitarian optimum is the endpoint farthest away from the mean of F (the barycenter of all homes).

b. Show that the majority relation coincides with the preferences of the median agent living at y^*. Exercise 4.7 generalizes this property.

*Exercise 4.6 Majority Relation under Single-Peaked Preferences

a. In example 4.6, show that the majority relation strictly prefers y to y' if and only if $|y - y^*| < |y' - y^*|$.

b. In the rest of the exercise, we only assume that individual preferences are single-peaked on [0, 1], as in the discussion following example 4.6. We denote by y^* the Condorcet winner.

Show that if $y < y' \leq y^*$, of $y^* \leq y' < y$, the majority relation R^m strictly prefers y' over y.

c. Check now that the majority relation is transitive *across its peak* y^*. Pick a, b, c such that $a < y^* < b < c$, and assume $aP^m b$. We know from question b that $bP^m c$. Show a $P^m c$. (*Hint:* An agent preferring a to b must have its peak to the left of b). Show similarly $aR^m b \Rightarrow aR^m c$.

Next pick a, b, c such that $b < a < y^* < c$, and assume $bP^m c$. Show $aP^m c$. (*Hint:* An agent preferring b to c must prefer a to c.) Conclude that the majority relation is single peaked, as claimed in section 4.4.

*Exercise 4.7 Majority Relation under Single-Dipped Preferences

We say that the preference relation R_i on [0, 1] is *single dipped* if there is an outcome x^i, the *dip,* such that R^i decreases on $[0, x^i]$ and increases on $[x^i, 1]$. In other words, x^i is the worst outcome for R_i and for all x, $x \neq x^i$, R_i prefers x to any outcome between x and x^i.

We fix a profile of single-dipped preferences on [0, 1], and we denote by y^* the median dip.

a. Show that the majority relation R^m is decreasing on $[0, y^*]$ and increasing on $[y^*, 1]$:

$$\{y < y' \leq y^* \text{ or } y^* \leq y' < y\} \Rightarrow yP^m y'$$

b. Show that R^m is transitive across its dip y^* (the argument is similar to that in question c of exercise 4.6.).

Exercise 4.8 Location of a Facility on a Network with a Loop

When the set of feasible locations of the facility is a one-dimensional network with a loop, the Condorcet winner often does not exist.

a. Consider the road network of example 3.8, depicted as figure 3.6. There are five agents living at A, B, C, D, and E respectively. If the "inner roads" to X are not feasible locations, it follows from example 4.6 that location C, the utilitarian optimum, is also the Condorcet winner. Assume now that the entire network on figure 3.6 is feasible to locate the facility.

Show that X, the utilitarian optimum is *not* a Condorcet winner, and neither is any other location.

b. Consider the road network of exercise 3.4, question a, depicted as figure 3.11. There are five agents living at A, B, C, D, and E respectively. Show similarly that no location of the facility is a Condorcet winner.

c. Consider the road network of exercise 3.4, question b, depicted as figure 3.12. There are nine agents, of whom two live at B, three at C, and four at D. Assume first that the direct road between C and B is closed. Show that A, the utilitarian optimum, is also the Condorcet winner.

Next show that when the road CB is open, the problem has no Condorcet winner.

*Exercise 4.9 Location of a Facility on a Star-Tree

The road network is a "star" as in exercise 3.6. Each outer location A_k is connected to the center O by a road of length d_k, and n_k agents live at A_k. See figure 3.13.

a. Assume that the facility is desirable, so an agent wants to minimize his travel cost to the facility. Show that the Condorcet winner coincides with the utilitarian optimum: it is the center O of the tree if none of the nodes A_k contains a strict majority of the total population; otherwise, it is this most populated location.

b. Now the facility is noxious and agents want it to be as far away as possible from where they live. Show that the Condorcet winner is A_{k^*}, the location farthest away from the center among those hosting a minority of the total population:

$$n_{k^*} < \frac{n}{2} \quad \text{and} \quad \left\{ \text{for all } k,\ n_k < \frac{n}{2} \Rightarrow d_{k^*} \geq d_k \right\}$$

(Note that the borderline case $n_1 = n_2 = n/2$ leads, as usual, to an "indecisive" majority relation.) Compare with the utilitarian optimum described in exercise 3.6.

Exercise 4.10 An Example with Intermediate Preferences

There are nine voters and three outcomes. Preferences are as follows:

Number of voters:	1	4	3	1
	b	a	b	a
	a	b	c	c
	c	c	a	b

a. Check that the Condorcet winner and Borda winners are different.

b. Show an ordering of the nine agents for which the profile has the intermediate preferences property.

Exercise 4.11 Voting over Surplus-Sharing and Taxation Methods

Using notations as in examples 4.7 and 4.8,

a. Give two examples of three person surplus-sharing problems with increasing claims $x_1 < x_2 < x_3$ and resources $t, t > x_1 + x_2 + x_3$, such that agent $2'$s shares are ordered respectively as follows:

$y_2(\text{ES}) > y_2(\text{PRO}) > y_2(\text{UG})$

$y_2(\text{ES}) > y_2(\text{UG}) > y_2(\text{PRO})$

b. Give two examples of three-person taxation problems with increasing taxable incomes $x_1 < x_2 < x_3$ and total aftertax income $t, t < x_1 + x_2 + x_3$, such that agent 2's aftertax incomes are ordered respectively as follows:

$y_2(\text{PRO}) > y_2(\text{UG}) > y_2(\text{UL})$

$y_2(\text{PRO}) > y_2(\text{UL}) > y_2(\text{UG})$

Note that example 4.8 contains a three-person example where $y_2(\text{PRO})$ is the smallest among these three shares.

Exercise 4.12 Voting over Taxation Methods

a. Consider the five voters profile of taxable incomes 10, 20, 50, 70, 80, and total aftertax income $t = 100$. Compute the aftertax distribution of incomes under the three basic methods PRO, UG, and UL and the majority relation.

b. Compute next the shares awarded by the Talmudic (T) and random priority (RP) methods defined in exercises 2.11 and 2.10 respectively. Check that the Talmudic method is the Condorcet winner and that RP is ranked third by the majority relation. Check that the Borda scores order the five methods in precisely the same way.

c. Consider the profile of taxable incomes 10, 15, 30, 40, 55, and aftertax income $t = 100$. Show that the Talmudic method is now the Condorcet loser as well as the Borda loser, with random priority ranked next to last in the majority and Borda relations.

*Exercise 4.13 Generalizing Example 4.7

We consider an arbitrary surplus-sharing problem with profile of claims x_i, $x_1 \leq x_2 \leq \cdots \leq x_n$ and resources t, $t \geq \sum_1^n x_i$.

a. Show that the profile of preferences over the three methods PRO, ES, and UG has the intermediate preferences property.

b. Show that the preferences of any agent over PRO, ES, UG are single-peaked when the three methods are so ordered.

Hint: By the inequalities proved in question a of exercise 2.5, it suffices to show that for any i, the inequalities $y_i(\mathrm{ES}) < y_i(\mathrm{PRO})$, $y_i(\mathrm{ES}) < y_i(\mathrm{UG})$, cannot both be true. Proceed by contradiction. The UG shares are $y_j(\mathrm{UG}) = \max\{x_j, \lambda\}$; show that $\lambda \leq t/n$ and that agent i for whom both inequalities hold has $x_i \leq \lambda$.

Exercise 4.14 Voting over the Commons

In the model of the commons of chapter 6, three solutions are compared: CEEI (competitive equilibrium with equal incomes), VP (virtual price), and RP (random priority). The users of the commons are identified by their willingness to pay, providing a natural ordering of N. Using the formulas of section 6.6, show that the preferences of the users over the three methods have the intermediate preference property. Show that in general, there is no ordering of CEEI, VP, and RP in which these preferences are single peaked.

Exercise 4.15 Counting Preference Relations

a. Given an ordering of the choice set A with cardinality p, show that there are 2^{p-1} different preference relations that are single peaked in this ordering.

b. Given a preference profile with the intermediate preferences property (with respect to some ordering of N), show that there are at most $[p(p-1)/2] + 1$ different preference relations in this profile.

Exercise 4.16 Proving the Claim in Example 4.8

Consider an arbitrary profile x_i, $x_1 \leq x_2 \leq \cdots \leq x_n$, of taxable incomes and total aftertax income t, $t \leq \sum_1^n x_i$.

a. Show that the profile of preferences over the three methods PRO, UG, UL has the intermediate preferences property. *Hint:* Check successively the property for any two of the three methods, and for both orderings of the two methods in each case.

b. Show that the IP property is maintained if we add the Talmudic (T) and random priority (RP) methods defined in exercises 2.11 and 2.10.

c. The goal of this question is to prove the IP property when the choice set contains all equal sacrifice methods listed in formula (6). We prove a slightly more general result. Consider two increasing utility functions u, v and the associated equal sacrifice after tax incomes y_i, z_i.

For simplicity, we assume first that all y_i, z_i are positive—system (9) in chapter 2—so that there are two positive numbers λ, μ such that

$$u(x_i) - u(y_i) = \lambda, \quad v(x_i) - v(z_i) = \mu \qquad \text{for all } i$$
$$\Leftrightarrow y_i = u^{-1}(u(x_i) - \lambda), \quad z = v^{-1}(v(x_i) - \mu) \qquad \text{for all } i$$

The desired IP property amounts to the fact that the two increasing functions, $f(x) = u^{-1}(u(x) - \lambda)$ and $g(x) = v^{-1}(v(x) - \mu)$, have the *single-crossing* property over positive numbers x: the graphs of these two functions cross at most once.

We assume now that the function v is more concave than u, namely $v(x) = w(u(x))$, all $x \geq 0$, for some increasing and concave function w. Check that this implies the single-crossing property. Check that if v is more concave than u, the IP property over u, v holds even if some agents receive a zero share—system (10) instead of (9) in chapter 2. Check finally that for any two utility functions in the family (6), one of them is more concave than the other.