CHAPTER

21

Social Choice Theory

21.A Introduction

In this chapter, we analyze the extent to which individual preferences can be aggregated into social preferences, or more directly into social decisions, in a "satisfactory" manner—that is, in a manner compatible with the fulfillment of a variety of desirable conditions.

Throughout the chapter, we contemplate a set of possible social alternatives and a population of individuals with well-defined preferences over these alternatives.

In Section 21.B, we start with the simplest case: that in which the set of alternatives has only two elements. There are then many satisfactory solutions to the aggregation problem. In our presentation, we focus on a detailed analysis of the properties of aggregation by means of majority voting.

In Section 21.C, we move to the case of many alternatives and the discussion takes a decidedly negative turn. We state and prove the celebrated *Arrow's impossibility theorem*. In essence, this theorem tells us that we cannot have everything: If we want our aggregation rule (which we call a *social welfare functional*) to be defined for any possible constellation of individual preferences, to always yield Pareto optimal decisions, and to satisfy the convenient, and key, property that social preferences over any two alternatives depend only on individual preferences over these alternatives (the *pairwise independence condition*), then we have a dilemma. Either we must give up the hope that social preferences could be rational in the sense introduced in Chapter 1 (i.e., that society behaves as an individual would) or we must accept dictatorship.

Section 21.D describes two ways out of the conclusion of the impossibility theorem. In one we allow for partial relaxations of the degree of rationality demanded of social preferences. In the other, we settle for aggregation rules that perform satisfactorily on restricted domains of individual preferences. In particular, we introduce the important notion of *single-peaked* preferences and, for populations with preferences in this class, we analyze the role of a *median voter* in the workings of pairwise majority voting as an aggregation method.

Section 21.E sets the aggregation problem more directly as one of aggregating individual preferences into social decisions. It introduces the concept of a *social choice* *function*, and proceeds to give a version of the impossibility result for the latter. Essentially, this result is obtained by replacing the pairwise independence condition (which is meaningless in the context of this section) by a *monotonicity* condition on the social choice function. This condition provides an important link to the incentive-based theory of Chapter 23.

General references and surveys for the topics of this chapter are Arrow (1963), Moulin (1988), and Sen (1970) and (1986).

21.B A Special Case: Social Preferences over Two Alternatives

We begin our analysis of social choice by considering the simplest possible case: that in which there are only two alternatives over which to decide. We call these alternative x and alternative y. Alternative x, for example, could be the "status-quo," and alternative y might be a particular public project whose implementation is being contemplated.

The data for our problem are the individual preferences of the members of society over the two alternatives. We assume that there is a number $I < \infty$ of individuals, or *agents*. The family of individual preferences between the two alternatives can be described by a profile

$$(\alpha_1, \dots, \alpha_I) \in \mathbb{R}^I,$$

where α_i takes the value 1, 0, or -1 according to whether agent i prefers alternative x to alternative y, is indifferent between them, or prefers alternative y to alternative x, respectively.[1]

Definition 21.B.1: A *social welfare functional* (or *social welfare aggregator*) is a rule $F(\alpha_1, \dots, \alpha_I)$ that assigns a social preference, that is, $F(\alpha_1, \dots, \alpha_I) \in \{-1, 0, 1\}$, to every possible profile of individual preferences $(\alpha_1, \dots, \alpha_I) \in \{-1, 0, 1\}^I$.

All the social welfare functionals to be considered respect individual preferences in the weak sense of Definition 21.B.2.

Definition 21.B.2: The social welfare functional $F(\alpha_1, \dots, \alpha_I)$ is *Paretian*, or has the *Pareto property*, if it respects unanimity of strict preference on the part of the agents, that is, if $F(1, \dots, 1) = 1$ and $F(-1, \dots, -1) = -1$.

Example 21.B.1: Paretian social welfare functionals between two alternatives abound. Let $(\beta_1, \dots, \beta_I) \in \mathbb{R}^I_+$ be a vector of nonnegative numbers, not all zero. Then we could define

$$F(\alpha_1, \dots, \alpha_I) = \text{sign} \sum_i \beta_i \alpha_i,$$

where, recall, for any $a \in \mathbb{R}$, sign a equals 1, 0, or -1 according to whether $a > 0$, $a = 0$, or $a < 0$, respectively.

An important particular case is *majority voting*, where we take $\beta_i = 1$ for every i. Then $F(\alpha_1, \dots, \alpha_I) = 1$ if and only if the number of agents that prefer alternative x to alternative y is larger than the number of agents that prefer y to x. Similarly, $F(\alpha_1, \dots, \alpha_I) = -1$ if and only if those that prefer y to x are more numerous than those that prefer x to y. Finally, in case of equality of these two numbers, we have $F(\alpha_1, \dots, \alpha_I) = 0$, that is, social indifference. ■

Example 21.B.2: *Dictatorship.* We say that a social welfare functional is *dictatorial* if there is an agent h, called a *dictator*, such that, for any profile $(\alpha_1, \dots, \alpha_I)$, $\alpha_h = 1$ implies $F(\alpha_1, \dots, \alpha_I) = 1$ and, similarly, $\alpha_h = -1$ implies $F(\alpha_1, \dots, \alpha_I) = -1$. That is, the strict preference of the dictator prevails as the social preference. A dictatorial social welfare functional is Paretian in the sense of Definition 21.B.2. For the social welfare functionals of Example 21.B.1, we have dictatorship whenever $\alpha_h > 0$ for some agent h and $\alpha_i = 0$ for $i \neq h$, since then $F(\alpha_1, \dots, \alpha_I) = \alpha_h$. ■

The majority voting social welfare functional plays a leading benchmark role in social choice theory. In addition to being Paretian it has three important properties, which we proceed to state formally. The first (symmetry among agents) says that the social welfare functional treats all agents on the same footing. The second (neutrality between alternatives) says that, similarly, the social welfare functional does not a priori distinguish either of the two alternatives. The third (positive responsiveness) says, more strongly than the Paretian property of Definition 21.B.2, that the social welfare functional is sensitive to individual preferences.

Definition 21.B.3: The social welfare functional $F(\alpha_1, \dots, \alpha_I)$ is *symmetric among agents* (or *anonymous*) if the names of the agents do not matter, that is, if a permutation of preferences across agents does not alter the social preference. Precisely, let $\pi: \{1, \dots, I\} \to \{1, \dots, I\}$ be an onto function (i.e., a function with the property that for any i there is h such that $\pi(h) = i$). Then for any profile $(\alpha_1, \dots, \alpha_I)$ we have $F(\alpha_1, \dots, \alpha_I) = F(\alpha_{\pi(1)}, \dots, \alpha_{\pi(I)})$.

Definition 21.B.4: The social welfare functional $F(\alpha_1, \dots, \alpha_I)$ is *neutral between alternatives* if $F(\alpha_1, \dots, \alpha_I) = -F(-\alpha_1, \dots, -\alpha_I)$ for every profile $(\alpha_1, \dots, \alpha_I)$, that is, if the social preference is reversed when we reverse the preferences of all agents.

Definition 21.B.5: The social welfare functional $F(\alpha_1, \dots, \alpha_I)$ is *positively responsive* if, whenever $(\alpha_1, \dots, \alpha_I) \geq (\alpha'_1, \dots, \alpha'_I)$, $(\alpha_1, \dots, \alpha_I) \neq (\alpha'_1, \dots, \alpha'_I)$, and $F(\alpha'_1, \dots, \alpha'_I) \geq 0$, we have $F(\alpha_1, \dots, \alpha_I) = +1$. That is, if x is socially preferred or indifferent to y and some agents raise their consideration of x, then x becomes socially preferred.

It is simple to verify that majority voting satisfies the three properties of symmetry among agents, neutrality between alternatives, and positive responsiveness (see Exercise 21.B.1). As it turns out, these properties entirely characterize majority voting. The result given in Proposition 21.B.1 is due to May (1952).

1. In the whole of this chapter we make the restriction that only the agents' rankings between the two alternatives matter for the social decision between them. In Section 21.C we will state formally the principle involved. Note, in particular, that this specification precludes the use of any "cardinal" or "intensity" information between the two alternatives because this intensity can only be calibrated (perhaps using lotteries) by appealing to some third alternative. A fortiori, the specification also precludes the comparison of feelings of pleasure or pain across individuals. In Chapter 22, we discuss in some detail matters pertaining to the issue of interpersonal comparability of utilities.

Proposition 21.B.1: (*May's Theorem*) A social welfare functional $F(\alpha_1, \ldots, \alpha_I)$ is a majority voting social welfare functional if and only if it is symmetric among agents, neutral between alternatives, and positive responsive.

Proof: We have already argued that majority voting satisfies the three properties. To establish sufficiency note first that the symmetry property among agents means that the social preference depends only on the total number of agents that prefer alternative x to y, the total number that are indifferent, and the total number that prefer y to x. Given $(\alpha_1, \ldots, \alpha_I)$, denote

$$n^+(\alpha_1, \ldots, \alpha_I) = \#\{i: \alpha_i = 1\}, \text{ and } n^-(\alpha_1, \ldots, \alpha_I) = \#\{i: \alpha_i = -1\}.^2$$

Then symmetry among agents allows us to express $F(\alpha_1, \ldots, \alpha_I)$ in the form

$$F(\alpha_1, \ldots, \alpha_I) = G(n^+(\alpha_1, \ldots, \alpha_I), n^-(\alpha_1, \ldots, \alpha_I)).$$

Now suppose that $(\alpha_1, \ldots, \alpha_I)$ is such that $n^+(\alpha_1, \ldots, \alpha_I) = n^-(\alpha_1, \ldots, \alpha_I)$. Then $n^+(-\alpha_1, \ldots, -\alpha_I) = n^-(\alpha_1, \ldots, \alpha_I) = n^+(\alpha_1, \ldots, \alpha_I) = n^-(-\alpha_1, \ldots, -\alpha_I)$, and so

$$\begin{aligned} F(\alpha_1, \ldots, \alpha_I) &= G(n^+(\alpha_1, \ldots, \alpha_I), n^-(\alpha_1, \ldots, \alpha_I)) \\ &= G(n^+(-\alpha_1, \ldots, -\alpha_I), n^-(-\alpha_1, \ldots, -\alpha_I)) \\ &= F(-\alpha_1, \ldots, -\alpha_I) \\ &= -F(\alpha_1, \ldots, \alpha_I). \end{aligned}$$

The last equality follows from the neutrality between alternatives. Since the only number that equals its negative is zero, we conclude that if $n^+(\alpha_1, \ldots, \alpha_I) = n^-(\alpha_1, \ldots, \alpha_I)$ then $F(\alpha_1, \ldots, \alpha_I) = 0$.

Suppose next that $n^+(\alpha_1, \ldots, \alpha_I) > n^-(\alpha_1, \ldots, \alpha_I)$. Denote $H = n^+(\alpha_1, \ldots, \alpha_I)$, $J = n^-(\alpha_1, \ldots, \alpha_I)$; then $J < H$. Say, without loss of generality, that $\alpha_i = 1$ for $i \le H$ and $\alpha_i \le 0$ for $i > H$. Consider a new profile $(\alpha'_1, \ldots, \alpha'_I)$ defined by $\alpha'_i = \alpha_i = 1$ for $i \le J < H$, $\alpha'_i = 0$ for $J < i \le H$, and $\alpha'_i = \alpha_i \le 0$ for $i > H$. Then $n^+(\alpha'_1, \ldots, \alpha'_I) = J$ and $n^-(\alpha'_1, \ldots, \alpha'_I) = n^-(\alpha_1, \ldots, \alpha_I) = J$. Hence $F(\alpha'_1, \ldots, \alpha'_I) = 0$. But by construction, the alternative x has lost strength in the new individual preference. Indeed, $(\alpha_1, \ldots, \alpha_I) \ge (\alpha'_1, \ldots, \alpha'_I)$ and $\alpha_{J+1} = 1 > 0 = \alpha'_{J+1}$. Therefore, by the positive responsiveness property, we must have $F(\alpha_1, \ldots, \alpha_I) = 1$.

In turn, if $n^-(\alpha_1, \ldots, \alpha_I) > n^+(\alpha_1, \ldots, \alpha_I)$ then $n^+(-\alpha_1, \ldots, -\alpha_I) > n^-(-\alpha_1, \ldots, -\alpha_I)$ and so $F(-\alpha_1, \ldots, -\alpha_I) = 1$. Therefore, by neutrality among alternatives:

$$F(\alpha_1, \ldots, \alpha_I) = -F(-\alpha_1, \ldots, -\alpha_I) = -1.$$

We conclude that $F(\alpha_1, \ldots, \alpha_I)$ is indeed a majority voting social welfare functional. ∎

In Exercise 21.B.2, you are asked to find examples different from majority voting that satisfy any two of the three properties of Proposition 21.B.1.

21.C The General Case: Arrow's Impossibility Theorem

We now proceed to study the problem of aggregating individual preferences over any number of alternatives. We denote the set of alternatives by X, and assume that

2. Recall the notation $\# A$ = cardinality of the set A = number of elements in the set A.

there are I agents, indexed by $i = 1, \ldots, I$. Every agent i has a rational preference relation $\succsim_i$ defined on X. The strict preference and the indifference relation derived from $\succsim_i$ are denoted by $\succ_i$ and $\sim_i$, respectively.[3] In addition, it will often be convenient to assume that no two distinct alternatives are indifferent in an individual preference relation $\succsim_i$. It is therefore important, for clarity of exposition, to have a symbol for the set of all possible rational preference relations on X and for the set of all possible preference relations on X having the property that no two distinct alternatives are indifferent. We denote these sets, respectively, by $\mathscr{R}$ and $\mathscr{P}$. Observe that $\mathscr{P} \subset \mathscr{R}$.[4]

In parallel to Section 21.B, we can define a social welfare functional as a rule that assigns social preferences to profiles of individual preferences $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I$. Definition 21.C.1 below generalizes Definition 21.B.1 in two respects: it allows for any number of alternatives and it permits the aggregation problem to be limited to some given domain $\mathscr{A} \subset \mathscr{R}^I$ of individual profiles. In this section, however, we focus on the largest domains, that is, $\mathscr{A} = \mathscr{R}^I$ and $\mathscr{A} = \mathscr{P}^I$.

Definition 21.C.1: A *social welfare functional* (or *social welfare aggregator*) defined on a given subset $\mathscr{A} \subset \mathscr{R}^I$ is a rule $F: \mathscr{A} \to \mathscr{R}$ that assigns a rational preference relation $F(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}$, interpreted as the social preference relation, to any profile of individual rational preference relations $(\succsim_1, \ldots, \succsim_I)$ in the admissible domain $\mathscr{A} \subset \mathscr{R}^I$.

Note that, as we did in Section, 21.B the problem of social aggregation is being posed as one in which individuals are described exclusively by their preference relations over alternatives.[5]

For any profile $(\succsim_1, \ldots, \succsim_I)$, we denote by $F_p(\succsim_1, \ldots, \succsim_I)$ the strict preference relation derived from $F(\succsim_1, \ldots, \succsim_I)$. That is, we let $x \, F_p(\succsim_1, \ldots, \succsim_I) \, y$ if $x \, F(\succsim_1, \ldots, \succsim_I) \, y$ holds but $y \, F(\succsim_1, \ldots, \succsim_I) \, x$ does not. We say then that "x is socially preferred to y." We read $x \, F(\succsim_1, \ldots, \succsim_I) \, y$ as "x is socially at least as good as y."

Definition 21.C.2. (which generalizes Definition 21.B.2) isolates the social welfare functionals that satisfy a minimal condition of respect for individual preferences.

3. Recall from Section 1.B that $\succ_i$ is formally defined by letting $x \succ_i y$ if $x \succsim_i y$ holds but $y \succsim_i x$ does not. That is, x is preferred to y if x is at least as good as y but y is not as good as x. Also, the indifference relation $\sim_i$ is defined by letting $x \sim_i y$ if $x \succsim_i y$ and $y \succsim_i x$. From Proposition 1.B.1 we know that if $\succsim_i$ is rational, that is, complete and transitive, then $\succ_i$ is irreflexive ($x \succ_i x$ cannot occur) and transitive ($x \succ_i y$ and $y \succ_i z$ implies $x \succ_i z$). Similarly, $\sim_i$ is reflexive ($x \sim_i x$ for all $x \in X$), transitive ($x \sim_i y$ and $y \sim_i z$ implies $x \sim_i z$) and symmetric ($x \sim_i y$ implies $y \sim_i x$).

4. Formally, the preference relation $\succsim_i$ belongs to $\mathscr{P}$ if it is reflexive ($x \succsim_i x$ for every $x \in X$), transitive ($x \succsim_i y$ and $y \succsim_i z$ implies $x \succsim_i z$) and *total* (if $x \neq y$ then either $x \succsim_i y$ or $y \succsim_i x$, but not both). Such preference relations are often referred to as *strict preferences* (although *strict-total preferences* would be less ambiguous) or even as *linear orders*, because these are the properties of the usual "larger than or equal to" order in the real line.

5. In particular, there are no individual utility levels and, therefore, there is no meaningful sense in which any conceivable information on individual utility levels could be compared and matched up. We refer again to Chapter 22 (especially Section 22.D) for an analysis of the problem that focuses on the information used in the aggregation process.

Definition 21.C.2: The social welfare functional $F: \mathscr{A} \to \mathscr{R}$ is *Paretian* if, for any pair of alternatives $\{x, y\} \subset X$ and any preference profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, we have that x is socially preferred to y, that is, $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$, whenever $x \succ_i y$ for every i.

In Example 21.C.1 we describe an interesting class of Paretian social welfare functionals.

Example 21.C.1: *The Borda Count.* Suppose that the number of alternatives is finite. Given a preference relation $\succsim_i \in \mathscr{R}$ we assign a number of points $c_i(x)$ to every alternative $x \in X$ as follows. Suppose for a moment that in the preference relation $\succsim_i$ no two alternatives are indifferent. Then we put $c_i(x) = n$ if x is the nth ranked alternative in the ordering of $\succsim_i$. If indifference is possible in $\succsim_i$ then $c_i(x)$ is the average rank of the alternatives indifferent to x.[6] Finally, for any profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I$ we determine a social ordering by adding up points. That is, we let $F(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}$ be the preference relation defined by $x\, F(\succsim_1, \ldots, \succsim_I)\, y$ if $\sum_i c_i(x) \le \sum_i c_i(y)$. This preference relation is complete and transitive [it is represented by the utility function $-c(x) = -\sum_i c_i(x)$]. Moreover, it is Paretian since if $x \succ_i y$ for every i then $c_i(x) < c_i(y)$ for every i, and so $\sum_i c_i(x) < \sum_i c_i(y)$. ■

We next state an important restriction on social welfare functionals first suggested by Arrow (1963). The restriction says that the social preferences between any two alternatives depend only on the individual preferences between the same two alternatives. There are three possible lines of justification for this assumption. The first is strictly normative and has considerable appeal: it argues that in settling on a social ranking between x and y, the presence or absence of alternatives other than x and y should not matter. They are irrelevant to the issue at hand. The second is one of practicality. The assumption enormously facilitates the task of making social decisions because it helps to separate problems. The determination of the social ranking on a subset of alternatives does not need any information on individual preferences over alternatives outside this subset. The third relates to incentives and belongs to the subject matter of Chapter 23 (see also Proposition 21.E.2). Pairwise independence is intimately connected with the issue of providing the right inducements for the truthful revelation of individual preferences.

Definition 21.C.3: The social welfare functional $F: \mathscr{A} \to \mathscr{R}$ defined on the domain $\mathscr{A}$ satisfies the *pairwise independence condition* (or the *independence of irrelevant alternatives condition*) if the social preference between any two alternatives $\{x, y\} \subset X$ depends only on the profile of individual preferences over the same alternatives. Formally[7], for any pair of alternatives $\{x, y\} \subset X$, and for any pair of preference profiles $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ and $(\succsim'_1, \ldots, \succsim'_I) \in \mathscr{A}$ with the property that, for every i,

$$x \succsim_i y \;\Leftrightarrow\; x \succsim'_i y \quad \text{and} \quad y \succsim_i x \;\Leftrightarrow\; y \succsim'_i x,$$

6. Thus if $X = \{x, y, z\}$ and $x \succ_i y \sim_i z$ then $c_i(x) = 1$, and $c_i(y) = c_i(z) = 2.5$.

7. The expressions that follow are a bit cumbersome, We emphasize therefore that they do nothing more than to capture formally the statement just made. An equivalent formulation would be: for any $\{x, y\} \subset X$, if $\succsim_i|\{x, y\} = \succsim'_i|\{x, y\}$ for all i, then $F(\succsim_1, \ldots, \succsim_I)|\{x, y\} = F(\succsim'_1, \ldots, \succsim'_I)|\{x, y\}$. Here $\succsim|\{x, y\}$ stands for the restriction of the preference ordering $\succsim$ to the set $\{x, y\}$.

we have that

$$x\, F(\succsim_1, \ldots, \succsim_I)\, y \quad\Leftrightarrow\quad x\, F(\succsim'_1, \ldots, \succsim'_I)\, y$$

and

$$y\, F(\succsim_1, \ldots, \succsim_I)\, x \quad\Leftrightarrow\quad y\, F(\succsim'_1, \ldots, \succsim'_I)\, x.$$

Example 21.C.1: *continued.* Alas, the Borda count does not satisfy the pairwise independence condition. The reason is simple: the rank of an alternative depends on the placement of *every* other alternative. Suppose, for example, that there are two agents and three alternatives $\{x, y, z\}$. For the preferences

$$x \succ_1 z \succ_1 y,$$
$$y \succ_2 x \succ_2 z$$

we have that x is socially preferred to y [indeed, $c(x) = 3$ and $c(y) = 4$]. But for the preferences

$$x \succ'_1 y \succ'_1 z,$$
$$y \succ'_2 z \succ'_2 x$$

we have that y is socially preferred to x [indeed, now $c(x) = 4$ and $c(y) = 3$]. Yet the relative ordering of x and y has not changed for either of the two agents.

For another illustration, this time with three agents and four alternatives $\{x, y, z, w\}$, consider

$$z \succ_1 x \succ_1 y \succ_1 w,$$
$$z \succ_2 x \succ_2 y \succ_2 w,$$
$$y \succ_3 z \succ_3 w \succ_3 x.$$

Here, y is socially preferred to x [$c(x) = 8$ and $c(y) = 7$]. But suppose now that alternatives z and w move to the bottom for all agents (which because of the Pareto property is a way of saying that the two alternatives are eliminated from the alternative set):

$$\begin{aligned} &x \succ'_1 y \succ'_1 z \succ'_1 w,\\ &x \succ'_2 y \succ'_2 z \succ'_2 w,\\ &y \succ'_3 x \succ'_3 z \succ'_3 w.\end{aligned} \tag{21.C.1}$$

Then x is socially preferred to y [$c(x) = 4$, $c(y) = 5$]. Thus the presence or absence of alternatives z and w matters to the social preference between x and y. Another modification would take alternative x to the bottom for agent 3:

$$x \succ''_1 y \succ''_1 z \succ''_1 w,$$
$$x \succ''_2 y \succ''_2 z \succ''_2 w,$$
$$y \succ''_3 z \succ''_3 w \succ''_3 x.$$

Now y is socially preferred to x [which, relative to the outcome with (21.C.1), is a nice result from the point of view of agent 3]. ■

The previous discussion of Example 21.C.1 teaches us that the pairwise independence condition is a substantial restriction. However, there is a way to proceed that will automatically guarantee that it is satisfied. It consists of determining the social preference between any given two alternatives by applying an aggregation rule that uses only the information about the ordering of *these two alternatives* in individual preferences. We saw in Section 21.B that, for any pair of alternatives, there are many such rules. Can we proceed in this pairwise fashion and still end up with social preferences that are rational, that is, complete and transitive? Example 21.C.2 shows that this turns out to be a real difficulty.

Example 21.C.2: *The Condorcet Paradox.*[8] Suppose that we were to try majority voting among any two alternatives (see Section 21.B for an analysis of majority voting). Does this determine a social welfare functional? We shall see in the next section that the answer is positive in some restricted domains $\mathscr{A} \subset \mathscr{R}^I$. But in general we run into the following problem, known as the Condorcet paradox. Let us have three alternatives $\{x, y, z\}$ and three agents. The preferences of the three agents are

$$x \succ_1 y \succ_1 z,$$
$$z \succ_2 x \succ_2 y,$$
$$y \succ_3 z \succ_3 x.$$

Then pairwise majority voting tells us that x must be socially preferred to y (since x has a majority against y and, a fortiori, y does not have a majority against x). Similarly, y must be socially preferred to z (two voters prefer y to z) and z must be socially preferred to x (two voters prefer z to x). But this cyclic pattern violates the transitivity requirement on social preferences. ■

The next proposition is *Arrow's impossibility theorem*, the central result of this chapter. It essentially tells us that the Condorcet paradox is not due to any of the strong properties of majority voting (which, we may recall from Proposition 21.B.1, are symmetry among agents, neutrality between alternatives, and positive responsiveness). The paradox goes to the heart of the matter: with pairwise independence there is no social welfare functonal defined on $\mathscr{R}^I$ that satisfies a minimal form of symmetry among agents (no dictatorship) and a minimal form of positive responsiveness (the Pareto property).

Proposition 21.C.1: (*Arrow's Impossibility Theorem*) Suppose that the number of alternatives is at least three and that the domain of admissible individual profiles, denoted $\mathscr{A}$, is either $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$. Then every social welfare functional $F: \mathscr{A} \to \mathscr{R}$ that is Paretian and satisfies the pairwise independence condition is *dictatorial* in the following sense: There is an agent h such that, for any $\{x, y\} \subset X$ and any profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, we have that x is socially preferred to y, that is, $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$, whenever $x \succ_h y$.

Proof: We present here the classical proof of this result. For another approach to the demonstration we refer to Section 22.D.

It is convenient from now on to view I not only as the number but also as the *set* of agents. For the entire proof we refer to a fixed social welfare functional $F: \mathscr{A} \to \mathscr{R}$ satisfying the Pareto and the pairwise independence conditions. We begin with some definitions. In what follows, when we refer to pairs of alternatives we always mean distinct alternatives.

8. This example was already discussed in Section 1.B.

Definition 21.C.4: Given $F(\cdot)$, we say that a subset of agents $S \subset I$ is:

(i) *Decisive for x over y* if whenever every agent in S prefers x to y *and* every agent not in S prefers y to x, x is socially preferred to y.
(ii) *Decisive* if, for any pair $\{x, y\} \subset X$, S is decisive for x over y.
(iii) *Completely decisive for x over y* if whenever every agent in S prefers x to y, x is socially preferred to y.

The proof will proceed by a detailed investigation of the structure of the family of decisive sets. We do this in a number of small steps. Steps 1 to 3 show that if a subset of agents is decisive for some pair of alternatives then it is decisive for all pairs. Steps 4 to 6 establish some algebraic properties of the family of decisive sets. Steps 7 and 8 use these to show that there is a smallest decisive set formed by a single agent. Steps 9 and 10 prove that this agent is a dictator.

Step 1: If for some $\{x, y\} \subset X$, $S \subset I$ is decisive for x over y, then, for any alternative $z \neq x$, S is decisive for x over z. Similarly, for any $z \neq y$, S is decisive for z over y.

We show that if S is decisive for x over y then it is decisive for x over any $z \neq x$. The reasoning for z over y is identical (you are asked to carry it out in Exercise 21.C.1).

If $z = y$ there is nothing to prove. So we assume that $z \neq y$. Consider a profile of preferences $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ where

$$x \succ_i y \succ_i z \quad \text{for every } i \in S$$

and

$$y \succ_i z \succ_i x \quad \text{for every } i \in I \backslash S.$$

Then, because S is decisive for x over y, we have that x is socially preferred to y, that is, $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$. In addition, since $y \succsim_i z$ for every $i \in I$, and $F(\cdot)$ satisfies the Pareto property it follows that $y\, F_p(\succsim_1, \ldots, \succsim_I)\, z$. Therefore, by the transitivity of the social preference relation, we conclude that $x\, F_p(\succsim_1, \ldots, \succsim_I)\, z$. By the pairwise independence condition, it follows that x is socially preferred to z whenever every agent in S prefers x to z and every agent not in S prefers z to x. That is, S is decisive for x over z.

Step 2: If for some $\{x, y\} \subset X$, $S \subset I$ is decisive for x over y and z is a third alternative, then S is decisive for z over w and for w over z, where $w \in X$ is any alternative distinct from z.

By step 1, S is decisive for x over z and for z over y. But then, applying step 1 again, this time to the pair $\{x, z\}$ and the alternative w, we conclude that S is decisive for w over z. Similarly, applying step 1 to $\{z, y\}$ and w, we conclude that S is decisive for z over w.

Step 3: If for some $\{x, y\} \subset X$, $S \subset I$ is decisive for x over y, then S is decisive.

This is an immediate consequence of step 2 and the fact that there is some alternative $z \in X$ distinct from x or y. Indeed, take any pair $\{v, w\}$. If $v = z$ or $w = z$, then step 2 implies the result directly. If $v \neq z$ and $w \neq z$, we apply step 2 to conclude that S is decisive for z over w, and then step 1 [applied to the pair $\{z, w\}$] to conclude that S is decisive for v over w.

Step 4: If $S \subset I$ and $T \subset I$ are decisive, then $S \cap T$ is decisive.

Take any triple of distinct alternatives $\{x, y, z\} \subset X$ and consider a profile of preferences $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ where

$$\begin{aligned} z \succ_i y \succ_i x \quad & \text{for every } i \in S \backslash (S \cap T), \\ x \succ_i z \succ_i y \quad & \text{for every } i \in S \cap T, \\ y \succ_i x \succ_i z \quad & \text{for every } i \in T \backslash (S \cap T), \\ y \succ_i z \succ_i x \quad & \text{for every } i \in I \backslash (S \cup T). \end{aligned}$$

Then $z\, F_p(\succsim_1, \ldots, \succsim_I)\, y$ because $S\,(=[S\backslash(S \cap T)] \cup (S \cap T))$ is a decisive set. Similarly, $x\, F_p(\succsim_1, \ldots, \succsim_I)\, z$ because T is a decisive set. Therefore, by the transitivity of the social preference, we have that $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$. It follows by the pairwise independence condition that $S \cap T$ is decisive for x over y, and so, by step 3, that $S \cap T$ is a decisive set.

Step 5: For any $S \subset I$, we have that either S or its complement, $I \backslash S \subset I$, is decisive.

Take any triple of distinct alternatives $\{x, y, z\} \subset X$ and consider a profile of preferences $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ where

$$\begin{aligned} x \succ_i z \succ_i y \quad & \text{for every } i \in S \\ y \succ_i x \succ_i z \quad & \text{for every } i \in I \backslash S. \end{aligned}$$

Then there are two possibilities: either $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$, in which case, by the pairwise independence condition, S is decisive for x over y (hence, by step 3, decisive), or $y\, F(\succsim_1, \ldots, \succsim_I)\, x$. Because, by the Paretian condition, we have $x\, F_p(\succsim_1, \ldots, \succsim_I)\, z$, the transitivity of the social preference relation yields that $y\, F_p(\succsim_1, \ldots, \succsim_I)\, z$ in this case. But then, using the pairwise independence condition again, we conclude that $I \backslash S$ is decisive for y over z (hence, by step 3, decisive).

Step 6: If $S \subset I$ is decisive and $S \subset T$, then T is also decisive.

Because of the Pareto property the empty set of agents cannot be decisive (indeed, if no agent prefers x over y and every agent prefers y over x, then x is not socially preferred to y). Therefore $I \backslash T$ cannot be decisive because otherwise, by step 4, $S \cap (I \backslash T) = \varnothing$ would be decisive. Hence, by step 5, T is decisive.

Step 7: If $S \subset I$ is decisive and it includes more than one agent, then there is a strict subset $S' \subset S$, $S' \neq S$, such that S' is decisive.

Take any $h \in S$. If $S \backslash \{h\}$ is decisive, then we are done. Suppose, therefore, that $S \backslash \{h\}$ is not decisive. Then, by step 5, $I \backslash (S \backslash \{h\}) = (I \backslash S) \cup \{h\}$ is decisive. It follows, by step 4, that $\{h\} = S \cap [(I \backslash S) \cup \{h\}]$ is also decisive. Thus, we are again done since, by assumption, $\{h\}$ is a strict subset of S.

Step 8: There is an $h \in I$ such that $S = \{h\}$ is decisive.

This follows by iterating step 7 and taking into account, first, that the set of agents I is finite and, second, that, by the Pareto property, the set I of all agents is decisive.

Step 9: If $S \subset I$ is decisive then, for any $\{x, y\} \subset X$, S is completely decisive for x over y.

We want to prove that, for any $T \subset I \backslash S$, x is socially preferred to y whenever every agent in S prefers x to y, every agent in T regards x to be at least as good as y, and every other agent prefers y to x. To verify this property, take a third alternative $z \in X$, distinct from x and y. By the pairwise independence condition it suffices to consider a profile of preferences $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ where

$$\begin{aligned} x \succ_i z \succ_i y \quad & \text{for every } i \in S, \\ x \succ_i y \succ_i z \quad & \text{for every } i \in T, \\ y \succ_i z \succ_i x \quad & \text{for every } i \in I \backslash (S \cup T). \end{aligned}$$

Then $x\, F_p(\succsim_1, \ldots, \succsim_I)\, z$ because, by step 6, $S \cup T$ is decisive, and $z\, F_p(\succsim_1, \ldots, \succsim_I)\, y$ because S is decisive. Therefore, by the transitivity of social preference, we have that $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$, as we wanted to show.

Step 10: If, for some $h \in I$, $S = \{h\}$ is decisive, then h is a dictator.

If $\{h\}$ is decisive then, by step 9, $\{h\}$ is completely decisive for any x over any y. That is, if the profile $(\succsim_1, \ldots, \succsim_I)$ is such that $x \succ_h y$, then $x\, F_p(\succsim_1, \ldots, \succsim_I)\, y$. But this is precisely what is meant by $h \in I$ being a dictator.

The combination of steps 8 and 10 completes the proof of Proposition 21.C.1. ■

21.D Some Possibility Results: Restricted Domains

The result of Arrow's impossibility theorem is somewhat disturbing, but it would be too facile to conclude from it that "democracy is impossible." What it shows is something else—that we should not expect a collectivity of individuals to behave with the kind of coherence that we may hope from an individual.

It is important to observe, however, that in practice collective judgments are made and decisions are taken. What Arrow's theorem does tell us, in essence, is that the institutional detail and procedures of the political process cannot be neglected. Suppose, for example, that the decision among three alternatives $\{x, y, z\}$ is made by first choosing between x and y by majority voting, and then voting again to choose between the winner and the third alternative z. This will produce an outcome, but the outcome may depend on how the agenda is set—that is, on which alternative is taken up first and which is left for the last. [Thus, if preferences are as in the Condorcet paradox (Example 21.C.2) then the last alternative, whichever it is, will always be the survivor.] This relevance of procedures and rules to social aggregation has far-reaching implications. They have been taken up and much emphasized in modern political science; see, for example, Austen-Smith and Banks (1996) or Shepsle and Boncheck (1995).

In this section, we remain modest and retain the basic framework. We explore to what extent we can escape the dictatorship conclusion if we relax some of the demands imposed by Arrow's theorem. We will investigate two weakenings. In the first, we relax the rationality requirements made on aggregate preferences. In the second, we pose the aggregation question in a restricted domain. In particular, we will consider a restriction—*single-peakedness*—that has been found to be significant and useful in applications.

Less Than Full Social Rationality

Suppose that we keep the Paretian and pairwise independence conditions but permit the social preferences to be less than fully rational. Two weakenings of the rationality of preferences are captured in Definition 21.D.1.

Definition 21.D.1: Suppose that the preference relation $\succsim$ on X is reflexive and complete. We say then that:

(i) $\succsim$ is *quasitransitive* if the strict preference $\succ$ induced by $\succsim$ (i.e. $x \succ y \Leftrightarrow x \succsim y$ but not $y \succsim x$) is transitive.
(ii) $\succsim$ is *acyclic* if $\succsim$ has a maximal element in every finite subset $X' \subset X$, that is, $\{x \in X' : x \succsim y \text{ for all } y \in X'\} \neq \varnothing$.

A quasitransitive preference relation is acyclic, but the converse may not hold. Also, a rational preference relation is quasitransitive, but, again, the converse may not hold.[9] Thus the weaker condition is acyclicity. Yet acyclicity is not a drastic weakening of rationality: Note, for example, that the social orderings of the Condorcet paradox (Example 21.C.2) also violate acyclicity. (For more on acylicity see Exercise 21.D.1.)

We will not discuss in detail the possibilities opened to us by these weakenings of social rationality. There are some but they are not very substantial. We refer to Sen (1970) for a detailed exposition. The next two examples are illustrative.

Example 21.D.1: *Oligarchy.* Let I be the set of agents, and let $S \subset I$ be a given subset of agents to be called an *oligarchy* (the possibilities $S = \{h\}$ or $S = I$ are permitted). Given any profile $(\succsim_1, \dots, \succsim_I) \in \mathscr{A}^I$, the social preferences are formed as follows: For any $x, y \in X$, we say that x *is socially at least as good as* y [written $x\, F(\succsim_1, \dots, \succsim_I)\, y$] *if there is at least one* $h \in S$ *that has* $x \succsim_h y$. Hence, x is socially *preferred* to y if and only if *every* member of the oligarchy prefers x to y. In Exercise 21.D.2 you should verify that this social preference relation is quasitransitive but not transitive (because social indifference fails to be transitive). This is the only condition of Arrow's impossibility theorem that is violated (the Paretian condition and pairwise independence conditions are clearly satisfied). Nonetheless, this is scarcely a satisfactory solution to the social aggregation problem, as the aggregator has become very sluggish. At one extreme, if the oligarchy is a single agent then we have a dictatorship. At the other, if the oligarchy is the entire population then society is able to express strict preference only if there is complete unanimity among its members. ■

Example 21.D.2: *Vetoers.* Suppose there are two agents and three alternatives $\{x, y, z\}$. Then given any profile of preferences $(\succsim_1, \succsim_2)$, we let the social preferences coincide with the preferences of agent 1 with one qualification: agent 2 can veto the possibility that alternative x be socially preferred to y. Specifically, if $y \succ_2 x$ then y is socially at least as good as x. Summarizing, for any two alternatives $\{v, w\} \subset \{x, y, z\}$ we have that v is socially at least as good as w if either $v \succsim_1 w$, or $v = y$, $w = x$ and $v \succ_2 w$. In Exercise 21.D.3 you should verify that the social preferences so defined are acyclic but not necessarily quasitransitive. ■

9. Suppose that $\succsim$ is quasitransitive. Assume for a moment that it is not acyclic. Then there is some finite set $X' \subset X$ without a maximal element for $\succsim$. That is, for every $x \in X'$ there is some $y \in X'$ such that $y \succ x$ (i.e., such that $y \succsim x$ but not $x \succsim y$). Thus, for any integer M we can find a chain $x^1 \succ x^2 \succ \cdots \succ x^M$, where $x^m \in X'$ for every $m = 1, \dots, M$. If M is larger than the number of alternatives in X', then there must be some repetition in this chain. Say that $x^{m'} = x^m$ for $m > m'$. By quasitransitivity, $x^{m'} \succ x^m = x^{m'}$, which is impossible because $\succ$ is irreflexive by definition. Hence, $\succsim$ must be acyclic. An example of an acyclic but not quasitransitive relation will be given in Example 21.D.2. The relation $\succ$ derived from a rational preference relation $\succsim$ is transitive (Proposition 1.B.1). An example of a quasitransitive, but not rational, preference relation is given in Example 21.D.1.

Single-Peaked Preferences

We proceed now to present the most important class of restricted domain conditions: single-peakedness. We will then see that, in this restricted domain, nondictatorial aggregation is possible. In fact, with a small qualification, we will see that on this domain pairwise majority voting gives rise on this domain to a social welfare functional.

Definition 21.D.2: A binary relation $\geq$ on the set of alternatives X is a *linear order* on X if it is *reflexive* (i.e., $x \geq x$ for every $x \in X$), transitive (i.e., $x \geq y$ and $y \geq z$ implies $x \geq z$) and *total* (i.e., for any distinct $x, y \in X$, we have that either $x \geq y$ or $y \geq x$, but not both).

Example 21.D.3: The simplest example of a linear order occurs when X is a subset of the real line, $X \subset \mathbb{R}$, and $\geq$ is the natural "greater than or equal to" order of the real numbers. ■

Definition 21.D.3: The rational preference relation $\succsim$ is *single peaked* with respect to the linear order $\geq$ on X if there is an alternative $x \in X$ with the property that $\succsim$ is increasing with respect to $\geq$ on $\{y \in X : x \geq y\}$ and decreasing with respect to $\geq$ on $\{y \in X : y \geq x\}$. That is,

$$\text{If } x \geq z > y \quad \text{then} \quad z \succ y$$

and

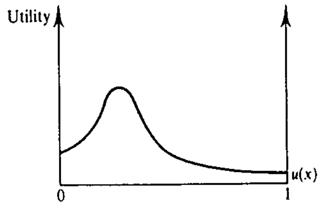
$$\text{If } y > z \geq x \quad \text{then} \quad z \succ y.$$

In words: There is an alternative x that represents a peak of satisfaction and, moreover, satisfaction increases as we approach this peak (so that, in particular, there cannot be any other peak of satisfaction).

Example 21.D.4: Suppose that $X = [a, b] \subset \mathbb{R}$ and $\geq$ is the "greater than or equal to" ordering of the real numbers. Then a continuous preference relation $\succsim$ on X is single peaked with respect to $\geq$ if and only if it is *strictly convex*, that is, if and only if, for every $w \in X$, we have $\alpha y + (1 - \alpha) z \succ w$ whenever $y \succsim w$, $z \succsim w$, $y \neq z$, and $\alpha \in (0,1)$. (Recall Definition 3.B.5 and also that, as a matter of definition, preference relations generated from strictly quasiconcave utility functions are strictly convex.) This fact accounts to a large extent for the importance of single-peakedness in economic applications. The sufficiency of strict convexity is actually quite simple to verify. (You are asked to prove necessity in Exercise 21.D.4.) Indeed, suppose that x is a maximal element for $\succsim$, and that, say, $x > z > y$. Then $x \succsim y$, $y \succsim y$, $x \neq y$, and $z = \alpha x + (1 - \alpha) y$ for some $\alpha \in (0, 1)$. Thus, $z \succ y$ by strict convexity. In Figures 21.D.1 and 21.D.2, we depict utility functions for two preference relations on $X = [0, 1]$. The preference relation in Figure 21.D.1 is single peaked with respect to $\geq$, but that in Figure 21.D.2 is not. ■

Definition 21.D.4: Given a linear order $\geq$ on X, we denote by $\mathscr{R}_{\geq} \subset \mathscr{R}$ the collection of all rational preference relations that are single peaked with respect to $\geq$.

Given a linear order $\geq$ and a set of agents I, from now on we consider the



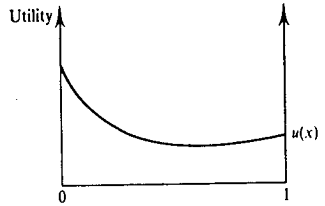


Figure 21.D.1 (left) Preferences are single peaked with respect to $\geq$.

Figure 21.D.2 (right) Preferences are not single peaked with respect to $\geq$.

restricted domain of preferences $\mathscr{R}^I_{\geq}$. This amounts to the requirement that all individuals have single-peaked preferences with respect to the *same* linear order $\geq$.

Suppose that on the domain $\mathscr{R}^I_{\geq}$ we define social preferences by means of pairwise majority voting (as introduced in Example 21.B.1). That is, given a profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I_{\geq}$ and any pair $\{x, y\} \subset X$, we put $x\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,y$, to be read as "$x$ is socially at least as good as y", if the number of agents that strictly prefer x to y is larger or equal to the number of agents that strictly prefer y to x, that is, if $\#\{i \in I : x \succ_i y\} \geq \#\{i \in I : y \succ_i x\}$.

Note that, from the definition, it follows that for any pair $\{x, y\}$ we must have either $x\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,y$ or $y\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,x$. Thus, pairwise majority voting induces a complete social preference relation (this holds on any possible domain of preferences).

In Exercise 21.D.5 you are asked to show in a direct manner that the preferences of the Condorcet paradox (Example 21.C.2) are not single peaked with respect to any possible linear order on the alternatives. In fact, they cannot be because, as we now show, with single-peaked preferences we are always assured that the social preferences induced by pairwise majority voting have maximal elements, that is, that there are alternatives that cannot be defeated by any other alternatives under majority voting.

Let $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I_{\geq}$ be a fixed profile of preferences. For every $i \in I$ we denote by $x_i \in X$ the maximal alternative for $\succsim_i$ (we will say that x_i is "i's peak").

Definition 21.D.5: Agent $h \in I$ is a *median agent for the profile* $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I_{\geq}$ if

$$\#\{i \in I : x_i \geq x_h\} \geq \frac{I}{2} \quad \text{and} \quad \#\{i \in I : x_h \geq x_i\} \geq \frac{I}{2}.$$

A median agent always exists. The determination of a median agent is illustrated in Figure 21.D.3.

If there are no ties in peaks and $\# I$ is odd, then Definition 21.D.5 simply says that a number $(I - 1)/2$ of the agents have peaks strictly smaller than x_h and another number $(I - 1)/2$ strictly larger. In this case the median agent is unique.

Proposition 21.D.1: Suppose that $\geq$ is a linear order on X and consider a profile of preferences $(\succsim_1, \ldots, \succsim_I)$ where, for every i, $\succsim_i$ is single peaked with respect to $\geq$. Let $h \in I$ be a median agent. Then $x_h\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,y$ for every $y \in X$. That is, the peak x_h of the median agent cannot be defeated by majority voting by any other alternative. Any alternative having this property is called a *Condorcet winner*. Therefore, a Condorcet winner exists whenever the preferences of all agents are singlepeaked with respect to the same linear order.

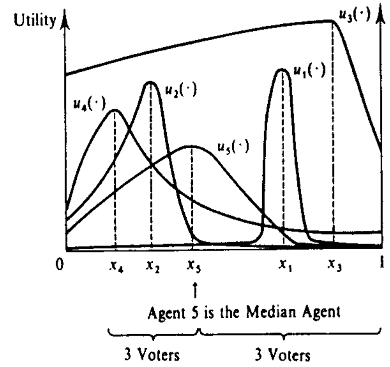


Figure 21.D.3 Determination of a median for a single-peaked family.

Proof: Take any $y \in X$ and suppose that $x_h > y$ (the argument is the same for $y > x_h$). We need to show that y does not defeat x, that is, that

$$\#\{i \in I : x_h \succ_i y\} \geq \#\{i \in I : y \succ_i x_h\}.$$

Consider the set of agents $S \subset I$ that have peaks *larger than or equal* to x_h, that is, $S = \{i \in I : x_i \geq x_h\}$. Then $x_i \geq x_h > y$ for every $i \in S$. Hence, by single-peakedness of $\succsim_i$ with respect to $\geq$, we get $x_h \succ_i y$ for every $i \in S$. On the other hand, because agent h is a median agent we have that $\# S \geq I/2$ and so $\#\{i \in I : y \succ_i x_h\} \leq \#(I \setminus S) \leq I/2 \leq \# S \leq \#\{i \in I : x_h \succ_i y\}$. ■

Proposition 21.D.1 guarantees that the preference relation $\hat{F}(\succsim_1, \ldots, \succsim_I)$ is acyclic. It may, however, not be transitive. In Exercise 21.D.6 you are asked to find an example of nontransitivity. Transitivity obtains in the special case where I is odd and, for every i, the preference relation $\succsim_i$ belongs to the class $\mathscr{P}^I_{\geq} \subset \mathscr{R}^I_{\geq}$ formed by the rational preference relations $\succsim$ that are single peaked with respect to $\geq$ and have the property that no two distinct alternatives are indifferent for $\succsim$. Note that, if I is odd and preferences are in this class, then, for any pair of alternatives, there is always a strict majority for one of them against the other. Hence, in this case, a Condorcet winner necessarily defeats any other alternative.

Proposition 21.D.2: Suppose that I is odd and that $\geq$ is a linear order on X. Then pairwise majority voting generates a well-defined social welfare functional $F : \mathscr{P}^I_{\geq} \to \mathscr{R}$. That is, on the domain of preferences that are single-peaked with respect to $\geq$ and, moreover, have the property that no two distinct alternatives are indifferent, we can conclude that the social relation $\hat{F}(\succsim_1, \ldots, \succsim_I)$ generated by pairwise majority voting is complete and transitive.

Proof: We already know that $\hat{F}(\succsim_1, \ldots, \succsim_I)$ is complete. It remains to show that it is transitive. For this purpose, suppose that $x\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,y$ and $y\,\hat{F}(\succsim_1, \ldots, \succsim_I)\,z$. Under our assumptions (recall that I is odd and that no individual indifference is allowed) this means that x defeats y and y defeats z. Consider the set $X' = \{x, y, z\}$. If preferences are restricted to this set then, relative to X', preferences still belong to the class $\mathscr{P}^I_{\geq}$, and therefore there is an alternative in X' that is not defeated by any

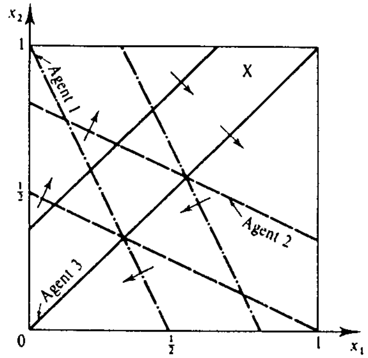


Figure 21.D.4
Indifference curves for the preferences of Example 21.D.5.

other alternative in X'. This alternative can be neither y (defeated by x) nor z (defeated by y). Hence, it has to be x and we conclude that $x\ \hat{F}(\succsim_1, \ldots, \succsim_I)\ z$, as required by transitivity. ∎

In applications, the linear order on alternatives arises typically as the natural order, as real numbers, of the values of a one-dimensional parameter. Then, as we have seen, single-peakedness follows from the strict quasiconcavity of utility functions, a restriction quite often satisfied in economics. It is an unfortunate fact that the power of quasiconcavity is confined to one-dimensional problems. We illustrate the issues involved in more general cases by discussing two examples.

Example 21.D.5: Suppose that the space of alternatives is the unit square, that is, $X = [0, 1]^2$. The generic entries of X are denoted $x = (x_1, x_2)$. There are three agents $I = \{1, 2, 3\}$. The preferences of the agents are expressed by the utility functions on X:

$$u_1(x_1, x_2) = -2x_1 - x_2,$$
$$u_2(x_1, x_2) = x_1 + 2x_2,$$
$$u_3(x_1, x_2) = x_1 - x_2.$$

These preferences are represented in Figure 21.D.4. Every utility function is linear and therefore preferences are convex (also, they have a single maximal element on X).[10] But, we will now argue that for every $x \in X$ there is a $y \in X$ preferred by two of the agents to x. To see this we take an arbitrary $x = (x_1, x_2) \in [0, 1]^2$ and distinguish three cases:

(i) If $x_1 = 0$, then $y = (\frac{1}{2}, x_2)$ is preferred by agents 2 and 3 to x.
(ii) If $x_2 = 1$, then $y = (x_1, \frac{1}{2})$ is preferred by agents 1 and 3 to x.
(iii) If $x_1 > 0$ and $x_2 < 1$, then $y = (x_1 - \varepsilon, x_2 + \varepsilon) \in [0, 1]^2$ with $\varepsilon > 0$, is preferred by agents 1 and 2 to x.

You should verify the claims made in (i), (ii), and (iii). ∎

The situation illustrated in Example 21.D.5 is not a peculiarity. The key property of the

10. The preferences of this example are not strictly convex. This is immaterial. Without changing the nature of the example we could modify them slightly so as to make the indifference curve map strictly convex.

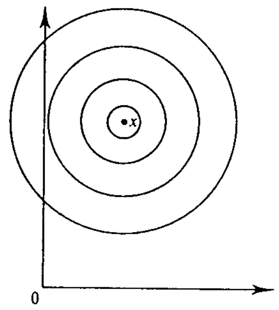


Figure 21.D.5 (left)
Euclidean preferences in $\mathbf{R}^2$.

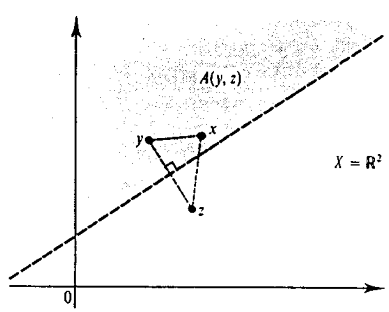


Figure 21.D.6 (right)
The region of Euclidean preferences that prefer y to z.

example is that the cone spanned by the nonnegative combinations of the gradient vectors of the three utility functions equals the entire $\mathbf{R}^2$ (see Figure 21.D.4). Exercises 21.D.7 and 21.D.8 provide further elaboration on this issue.

The reason why in two (or more) dimensions, quasiconcavity does not particularly help is that, in contrast with the one-dimensional case, there is no sensible way to assign a "median" to a set of points in the plane. This will become clear in the next, classical, Example 21.D.6 which we now describe.

Example 21.D.6: *Euclidean Preferences.* Suppose that the set of alternatives is $\mathbf{R}^n$. Agents have preferences $\succsim$ represented by utility functions of the form $u(y) = -\|y - x\|$, where x is a fixed alternative in $\mathbf{R}^n$. In words: x is the most preferred alternative for $\succsim$ and other alternatives are evaluated by how close they are to x in the Euclidean distance. The indifference curves of a typical consumer in $\mathbf{R}^2$ are pictured in Figure 21.D.5.

In the current example, the set $\mathbf{R}^n$ does double duty. On the one hand, it represents the set of alternatives. On the other, it also stands for the set of all possible preferences because every $x \in \mathbf{R}^n$ uniquely identifies the preferences that have x as a peak.[11]

Given two distinct alternatives $y, z \in \mathbf{R}^n$, an agent will prefer y to z if and only if his peak is closer to y than to z. Thus, the region of peaks associated with preferences that prefer y to z is

$$A(y, z) = \{x \in \mathbf{R}^n: \|x - y\| < \|x - z\|\}.$$

See Figure 21.D.6 for a representation. Geometrically, the boundary of $A(y, z)$ is the hyperplane perpendicular to the segment connecting y and z and passing through its midpoint.

We will consider the idealized limit situation where there is a continuum of agents with Euclidean preferences and the population is described by a density function $g(x)$ defined on $\mathbf{R}^n$, the set of possible peaks. Then given two distinct alternatives $y, z \in \mathbf{R}^n$, the fraction of the total population that prefers y to z, denoted $m_g(y, z)$, is simply the integral of $g(\cdot)$ over the region $A(y, z) \subset \mathbf{R}^n$.

When will there exist a Condorcet winner? Suppose there is an $x^* \in \mathbf{R}^n$ with the property that *any* hyperplane through x^* divides $\mathbf{R}^n$ into two half-spaces each having a total mass of $\frac{1}{2}$ according to the density $g(\cdot)$. This point could be called a *median* for the density $g(\cdot)$; it coincides with the usual concept of a median in the case $n = 1$. A median in this sense is a Condorcet winner. It cannot be defeated by any other alternative because if $y \neq x^*$ then $A(x^*, y)$ is larger than a half-space through x^* and, therefore, $m_g(x^*, y) \geq \frac{1}{2}$. Conversely, if x^*

11. For an example in the same spirit where the two roles are kept separate, see Grandmont (1978) and Exercise 21.D.9.

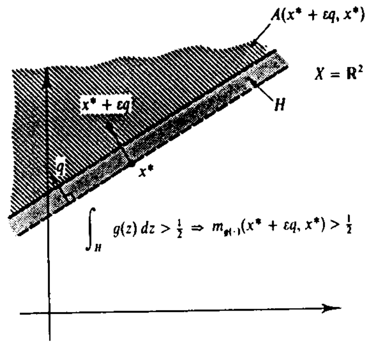


Figure 21.D.7
If x^* is not a median then it is not a Condorcet winner.

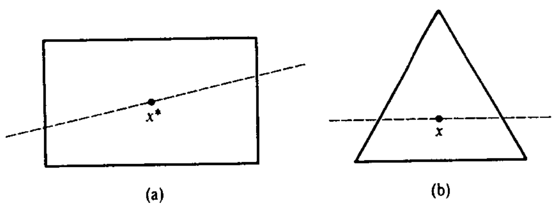


Figure 21.D.8
(a) Uniform distribution over a rectangle: The center point x^* is a median since every plane through x^* divides the rectangle into two figures of equal area.
(b) Uniform distribution over a triangle: There is no median.

is not a median then there is a direction $q \in \mathbf{R}^n$ such that the mass of the half-space $\{z \in \mathbf{R}^n : q \cdot z > q \cdot x^*\}$ is larger than $\frac{1}{2}$. Thus, by continuity, if $\varepsilon > 0$ is small then the mass of the translated half-space $A(x^* + \varepsilon q, x^*)$ is larger than $\frac{1}{2}$. Hence $x^* + \varepsilon q$ defeats x^*, and x^* cannot be a Condorcet winner. (See Figure 21.D.7.)

We have seen that a Condorcet winner exists if and only if there is a median for the density $g(\cdot)$. But for $n > 1$ the existence of a median imposes so many conditions (there are many half-spaces) that it becomes a knife-edge property. Figure 21.D.8 provides examples. In Figure 21.D.8(a), the density $g(\cdot)$ is the uniform density over a rectangle [a case first studied by Tullock (1967)]. The center of the rectangle is then, indeed, a median. But the rectangle is very special. The typical case is one of nonexistence. In Figure 21.D.8(b), the density $g(\cdot)$ is the uniform density over a triangle. Then no median exists: Through any point of a triangle we can draw a line that divides it into two regions of unequal area.[12] ■

12. See Caplin and Nalebuff (1988) for further analysis. They show that under a restriction on the density function (called "logarithmic concavity" and satisfied, in particular, for uniform densities over convex sets), there are always points ("generalized medians") in $\mathbf{R}^n$ with the property that any hyperplane through the point divides $\mathbf{R}^n$ into two regions, each of which has mass larger than $1/e$. This means that these points cannot be defeated by any other alternative *if* the majority required is not $\frac{1}{2}$ but any number larger than $1 - (1/e) > \frac{1}{2}$, 64% say. Of course, a 64% rule becomes less decisive than a 50% rule: There will now be many pairs of alternatives with the property that no member of the pair defeats the other.

21.E Social Choice Functions

The task we set ourselves to accomplish in the previous sections was how to aggregate profiles of individual preference relations into a coherent (i.e. rational) social preference order. Presumably, this social preference order is then used to make decisions. In this section we focus directly on social decisions and pose the aggregation question as one of analyzing how profiles of individual preferences turn into social decisions.

The main result we obtain again yields a dictatorship conclusion. The result amounts, in a sense, to a translation of the Arrow's impossibility theorem into the language of choice functions. It also offers a reinterpretation of the condition of pairwise independence, and provides a link towards the incentive-based analysis of Chapter 23.

As before, we have a set of alternatives X and a finite set of agents I. The set of preference relations $\succsim$ on X is denoted $\mathcal{R}$. We also designate by $\mathcal{P}$ the subset of $\mathcal{R}$ consisting of the preference relations $\succsim \in \mathcal{R}$ with the property that no two distinct alternatives are indifferent for $\succsim$.

Definition 21.E.1: Given any subset $\mathcal{A} \subset \mathcal{R}^I$, a *social choice function* $f: \mathcal{A} \to X$ defined on $\mathcal{A}$ assigns a chosen element $f(\succsim_1, \ldots, \succsim_I) \in X$ to every profile of individual preferences in $\mathcal{A}$.

The notion of social choice function embodies the requirement that the chosen set be single valued. We could argue that this is, after all, in the nature of what a choice is.[13] More restrictive is the fact that we do not allow for random choice.[14]

If X is finite, every social welfare functional $F(\cdot)$ on a domain $\mathcal{A}$ induces a natural social choice function by associating with each $(\succsim_1, \ldots, \succsim_I) \in \mathcal{A}$ a most preferred element in X for the social preference relation $F(\succsim_1, \ldots, \succsim_I)$. For example, if, as in Proposition 21.D.2, $\mathcal{A} \subset \mathcal{P}^I_{\geq}$ is a domain of single-peaked preferences, I is odd, and $F(\cdot)$ is the pairwise majority voting social welfare functional defined on $\mathcal{A}$, then for every $(\succsim_1, \ldots, \succsim_I)$ the choice $f(\succsim_1, \ldots, \succsim_I)$ is the Condorcet winner in X.

We now state and prove a result parallel to Arrow's impossibility theorem. Recall that for Arrow's theorem we had two conditions: the social welfare functional had to be Paretian and had to be pairwise independent. Here we require again two conditions: first, the social choice function must be, again, *(weakly) Paretian*; and, second, it should be *monotonic*. We define these concepts in Definitions 21.E.2 and 21.E.4, respectively.

13. Nevertheless, allowing for multivalued choice sets (that is, allowing there to be more than one acceptable social choice) is natural in some contexts, and certain assumptions on social choice may be more plausible in the multivalued case.

14. Note also the contrast between the definition of choice function here and the similar concept of choice rule in Section 1.C. There we contemplated the possibility of there being several budgets and of the choice depending on the budget at hand. Here the budget is fixed (it is always X) but the choice may depend on the profile of underlying individual preferences. Clearly we could, but will not, consider situations that encompass both cases. Another contrast with Section 1.C is that here we limit ourselves to single-valued choice.

Definition 21.E.2: The social choice function $f: \mathscr{A} \to X$ defined on $\mathscr{A} \subset \mathscr{R}^I$ is *weakly Paretian* if for any profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ the choice $f(\succsim_1, \ldots, \succsim_I) \in X$ is a weak Pareto optimum. That is, if for some pair $\{x, y\} \subset X$ we have that $x \succ_i y$ for every i, then $y \neq f(\succsim_1, \ldots, \succsim_I)$.

In order to define monotonicity we need a preliminary concept.

Definition 21.E.3: The alternative $x \in X$ *maintains its position from* the profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I$ to the profile $(\succsim'_1, \ldots, \succsim'_I) \in \mathscr{R}^I$ if

$$x \succsim_i y \quad \text{implies} \quad x \succsim'_i y$$

for every i and every $y \in X$.

In other words, x maintains its position from $(\succsim_1, \ldots, \succsim_I)$ to $(\succsim'_1, \ldots, \succsim'_I)$ if for every i the set of alternatives inferior (or indifferent) to x expands (or remains the same) in moving from $\succsim_i$ to $\succsim'_i$. That is,

$$L(x, \succsim_i) = \{y \in X : x \succsim_i y\} \subset L(x, \succsim'_i) = \{y \in X : x \succsim'_i y\}.$$

Note that the condition stated in Definition 21.E.3 imposes no restriction on how other alternatives different from x may change their mutual order in going from $\succsim_i$ to $\succsim'_i$.[15]

Definition 21.E.4: The social choice function $f: \mathscr{A} \to X$ defined on $\mathscr{A} \subset \mathscr{R}^I$ is *monotonic* if for any two profiles $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, $(\succsim'_1, \ldots, \succsim'_I) \in \mathscr{A}$ with the property that the chosen alternative $x = f(\succsim_1, \ldots, \succsim_I)$ maintains its position from $(\succsim_1, \ldots, \succsim_I)$ to $(\succsim'_1, \ldots, \succsim'_I)$, we have that $f(\succsim'_1, \ldots, \succsim'_I) = x$.

In words: The social choice function is monotonic if no alternative can be dropped from being chosen unless for some agent its disirability deteriorates.

Are there social choice functions that are weakly Paretian and monotonic? The answer is "yes." For example, in Exercise 21.E.1 you are asked to verify that the pairwise majority voting social decision function defined on a domain of single-peaked preferences is weakly Paretian and monotonic. But what if we have a universal domain (i.e., $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$)? A not very attractive class of social choice functions having the two properties in this domain are the *dictatorial* social choice functions.

Definition 21.E.5: An agent $h \in I$ is a *dictator* for the social choice function $f: \mathscr{A} \to X$ if, for every profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, $f(\succsim_1, \ldots, \succsim_I)$ is a most preferred alternative for $\succsim_h$ in X; that is,

$$f(\succsim_1, \ldots, \succsim_I) \in \{x \in X : x \succsim_h y \text{ for every } y \in X\}.$$

A social choice function that admits a dictator is called *dictatorial*.

In the domain $\mathscr{P}^I$, a dictatorial social choice function is weakly Paretian and monotonic. (This is clear enough, but at any rate you should verify it in Exercise 21.E.2, where you are also asked to discuss the case $\mathscr{A} = \mathscr{R}^I$.) Unfortunately, in the universal domain we cannot get anything better than the dictatorial social choice functions. The impossibility result of Proposition 21.E.1 establishes this.

15. As in Section 3.B, the sets $L(x, \succsim_i)$ are also referred to as *lower contour sets*.

Proposition 21.E.1: Suppose that the number of alternatives is at least three and that the domain of admissible preference profiles is either $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$. Then every weakly Paretian and monotonic social choice function $f: \mathscr{A} \to X$ is dictatorial.

Proof: The proof of the result will be obtained as a corollary of Arrow's impossibility theorem (Proposition 21.C.1). To this effect, we proceed to derive a social welfare functional $F(\cdot)$ that rationalizes $f(\succsim_1, \ldots, \succsim_I)$ for every profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$. We will then show that $F(\cdot)$ satisfies the assumptions of Arrow's theorem, hence yielding the dictatorship conclusion.

We begin with a useful definition.

Definition 21.E.6: Given a finite subset $X' \subset X$ and a profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I$, we say that the profile $(\succsim'_1, \ldots, \succsim'_I)$ *takes X' to the top from* $(\succsim_1, \ldots, \succsim_I)$ if, for every i,

$$x \succ'_i y \qquad \text{for } x \in X' \text{ and } y \notin X',$$
$$x \succsim_i y \Leftrightarrow x \succsim'_i y \qquad \text{for all } x, y \in X'.$$

In words: The preference relation $\succsim'_i$ is obtained from $\succsim_i$ by simply taking every alternative in X' to the top, while preserving the internal (weak or strict) ordering among these alternatives. The ordering among alternatives not in X' is arbitrary. For example, if $x \succ_i y \succ_i z \succ_i w$, then the preference relation $\succ'_i$ defined by $y \succ'_i w \succ'_i z \succ'_i x$ takes $\{y, w\}$ to the top from $\succsim_i$. Note also that if $(\succsim'_1, \ldots, \succsim'_I)$ takes X' to the top from $(\succsim_1, \ldots, \succsim_I)$, then every $x \in X'$ maintains its position in going from $(\succsim_1, \ldots, \succsim_I)$ to $(\succsim'_1, \ldots, \succsim'_I)$.

For the rest of the proof we proceed in steps:

Step 1: If both the profiles $(\succsim'_1, \ldots, \succsim'_I) \in \mathscr{A}$ and $(\succsim''_1, \ldots, \succsim''_I) \in \mathscr{A}$ take $X' \subset X$ to the top from $(\succsim_1, \ldots, \succsim_I)$, then $f(\succsim'_1, \ldots, \succsim'_I) = f(\succsim''_1, \ldots, \succsim''_I)$.

For every i and $x \in X'$ we have

$$\{y \in X : x \succsim'_i y\} = \{y \in X : x \succsim''_i y\} = \{y \in X : x \succsim_i y\} \cup X \backslash X'.$$

By the weak Pareto property, $f(\succsim'_1, \ldots, \succsim'_I) \in X'$. Thus, $f(\succsim'_1, \ldots, \succsim'_I) \in X'$ maintains its position in going from $(\succsim'_1, \ldots, \succsim'_I)$ to $(\succsim''_1, \ldots, \succsim''_I)$. Therefore, by the monotonicity of $f(\cdot)$, we conclude that $f(\succsim'_1, \ldots, \succsim'_I) = f(\succsim''_1, \ldots, \succsim''_I)$.

Step 2: Definition of $F(\succsim_1, \ldots, \succsim_I)$.

For every profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ we define a certain binary relation $F(\succsim_1, \ldots, \succsim_I)$ on X. Specifically, we let $x\, F(\succsim_1, \ldots, \succsim_I)\, y$, (read as "$x$ is socially at least as good as y") if $x = y$ or if $x = f(\succsim'_1, \ldots, \succsim'_I)$ when $(\succsim'_1, \ldots, \succsim'_I) \in \mathscr{A}$ is any profile that takes $\{x, y\} \subset X$ to the top from the profile $(\succsim_1, \ldots, \succsim_I)$. By step 1 this is well defined, that is, independent of the particular profile $(\succsim'_1, \ldots, \succsim'_I)$ chosen.

Step 3: For every profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, $F(\succsim_1, \ldots, \succsim_I)$ is a rational preference relation. Moreover, $F(\succsim_1, \ldots, \succsim_I) \in \mathscr{P}$; that is, no two distinct alternatives are socially indifferent.

Because $f(\cdot)$ is weakly Paretian, it follows that when $(\succsim'_1, \ldots, \succsim'_I)$ takes $\{x, y\}$ to the top from $(\succsim_1, \ldots, \succsim_I)$ we must have $f(\succsim'_1, \ldots, \succsim'_I) \in \{x, y\}$. Therefore, we conclude that either $x\, F(\succsim_1, \ldots, \succsim_I)\, y$ or $y\, F(\succsim_1, \ldots, \succsim_I)\, x$, *but*, because of step 1, *not both* (unless $x = y$). In particular, $F(\succsim_1, \ldots, \succsim_I)$ is complete.

To verify transitivity, suppose that $x\ F(\succsim_1,\dots,\succsim_I)\ y$ and $y\ F(\succsim_1,\dots,\succsim_I)\ z$. We can assume that the three alternatives $\{x, y, z\}$ are distinct. Let $(\succsim''_1,\dots,\succsim''_I) \in \mathscr{A}$ be a profile that takes $\{x, y, z\}$ to the top from $(\succsim_1,\dots,\succsim_I)$. Because $f(\cdot)$ is weakly Paretian, we have $f(\succsim''_1,\dots,\succsim''_I) \in \{x, y, z\}$.

Suppose that we had $y = f(\succsim''_1,\dots,\succsim''_I)$. Consider a profile $(\succsim'_1,\dots,\succsim'_I) \in \mathscr{A}$ that takes $\{x, y\}$ to the top from $(\succsim''_1,\dots,\succsim''_I)$. Since y maintains its position from $(\succsim''_1,\dots,\succsim''_I)$ to $(\succsim'_1,\dots,\succsim'_I)$, it follows from monotonicity that $f(\succsim'_1,\dots,\succsim'_I) = y$. But $(\succsim'_1,\dots,\succsim'_I)$ also takes $\{x, y\}$ to the top from $(\succsim_1,\dots,\succsim_I)$: the relative ordering of x and y, the two alternatives at the top, has not been altered in any individual preference in going from $(\succsim_1,\dots,\succsim_I)$ to $(\succsim'_1,\dots,\succsim'_I)$. Therefore we conclude that $y\ F(\succsim_1,\dots,\succsim_I)\ x$, which contradicts the assumption that $x\ F(\succsim_1,\dots,\succsim_I)\ y$, $x \neq y$. Hence, $y \neq f(\succsim''_1,\dots,\succsim''_I)$.

Similarly, we obtain $z \neq f(\succsim''_1,\dots,\succsim''_I)$. We only need to repeat the same argument using the pair $\{y, z\}$ (you are asked to do so in Exercise 21.E.3).

The only possibility left is $x = f(\succsim''_1,\dots,\succsim''_I)$. Thus, let $(\succsim'_1,\dots,\succsim'_I) \in \mathscr{A}$ take $\{x, z\}$ to the top from $(\succsim''_1,\dots,\succsim''_I)$. Since x maintains its position in going from $(\succsim''_1,\dots,\succsim''_I)$ to $(\succsim'_1,\dots,\succsim'_I)$, it follows that $x = f(\succsim'_1,\dots,\succsim'_I)$. But $(\succsim'_1,\dots,\succsim'_I)$ also takes $\{x, z\}$ to the top from $(\succsim_1,\dots,\succsim_I)$. Thus, $x\ F(\succsim_1,\dots,\succsim_I)\ z$, and transitivity is established.

Step 4: The social welfare functional $F: \mathscr{A} \to \mathscr{P}$ rationalizes $f: \mathscr{A} \to X$; that is, for every profile $(\succsim_1,\dots,\succsim_I) \in \mathscr{A}$, $f(\succsim_1,\dots,\succsim_I)$ is a most preferred alternative for $F(\succsim_1,\dots,\succsim_I)$ in X.

This is intuitive enough since $F(\cdot)$ has been constructed from $f(\cdot)$. Denote $x = f(\succsim_1,\dots,\succsim_I)$ and let $y \neq x$ be any other alternative. Consider a profile $(\succsim'_1,\dots,\succsim'_I) \in \mathscr{A}$ that takes $\{x, y\}$ to the top from $(\succsim_1,\dots,\succsim_I)$. Since x maintains position from $(\succsim_1,\dots,\succsim_I)$ to $(\succsim'_1,\dots,\succsim'_I)$, we have $x = f(\succsim'_1,\dots,\succsim'_I)$. Therefore, $x\ F(\succsim_1,\dots,\succsim_I)\ y$.

Step 5: The social welfare functional $F: \mathscr{A} \to \mathscr{P}$ is Paretian.

Clear if $x \succ_i y$ for every i then, by the Paretian property of $f(\cdot)$, we must have $x = f(\succsim'_1,\dots,\succsim'_I)$ whenever $(\succsim'_1,\dots,\succsim'_I)$ takes $\{x, y\}$ to the top from $(\succsim_1,\dots,\succsim_I)$. Hence $x\ F(\succsim_1,\dots,\succsim_I)\ y$, and by step 3 we conclude that $x\ F_p(\succsim_1,\dots,\succsim_I)\ y$.

Step 6: The social welfare functional $F: \mathscr{A} \to \mathscr{P}$ satisfies the pairwise independence condition.

This follows from step 1. Suppose that $(\succsim_1,\dots,\succsim_I) \in \mathscr{A}$ and $(\succsim'_1,\dots,\succsim'_I) \in \mathscr{A}$ have the same ordering of $\{x, y\}$ for every i (that is, for every i, $x \succsim_i y$ if and only if $x \succsim'_i y$). Suppose that $(\succsim''_1,\dots,\succsim''_I) \in \mathscr{A}$ takes $\{x, y\}$ to the top from $(\succsim_1,\dots,\succsim_I)$ and that, say, $x = f(\succsim''_1,\dots,\succsim''_I)$. Then $x\ F(\succsim_1,\dots,\succsim_I)\ y$. But $(\succsim''_1,\dots,\succsim''_I)$ also takes $\{x, y\}$ to the top from $(\succsim'_1,\dots,\succsim'_I)$. Hence, $x\ F(\succsim'_1,\dots,\succsim'_I) y$, as we wanted to prove.

Step 7: The social choice function $f: \mathscr{A} \to X$ is dictatorial.

By Arrow's theorem (Proposition 21.C.1) there is an agent $h \in I$ such that for every profile $(\succsim_1,\dots,\succsim_I) \in \mathscr{A}$ we have $x\ F_p(\succsim_1,\dots,\succsim_I) y$ whenever $x \succ_h y$. Therefore, $f(\succsim_1,\dots,\succsim_I)$ [which by step 4 is a most preferred alternative for

$F(\succsim_1,\dots,\succsim_I)$ in X] must also be a most preferred alternative for h; that is, $f(\succsim_1,\dots,\succsim_I) \succsim_h x$ for every $x \in X$. Hence agent h is a dictator. ■

Finally, we mention the following corollary (Proposition 21.E.2) to hint at the connection between Proposition 21.E.1 and the issue of incentives to truthful preference revelation, a topic that is studied extensively in Chapter 23.

Proposition 21.E.2: Suppose that the number of alternatives is at least three and that $f: \mathscr{P}^I \to X$ is a social choice function that is weakly Paretian and satisfies the following *no-incentive-to-misrepresent* condition:

$$f(\succsim_1,\dots,\succsim_{h-1},\succsim_h,\succsim_{h+1},\dots,\succsim_I) \succsim_h f(\succsim_1,\dots,\succsim_{h-1},\succsim'_h,\succsim_{h+1},\dots,\succsim_I)$$

for every agent h, every $\succsim'_h \in \mathscr{P}$, and every profile $(\succsim_1,\dots,\succsim_I) \in \mathscr{P}^I$. Then $f(\cdot)$ is dictatorial.

Proof: In view of Proposition 21.E.1 it suffices to show that $f: \mathscr{P}^I \to X$ must be monotonic.

Suppose that it is not. Then without loss of generality we can assume that, for some agent h, there are preferences $\succsim_i \in \mathscr{P}$ for agents $i \neq h$, and preferences $\succsim''_h$, $\succsim'''_h \in \mathscr{P}$ for agent h, such that, denoting

$$x = f(\succsim_1,\dots,\succsim_{h-1},\succsim''_h,\succsim_{h+1},\dots,\succsim_I)$$

and

$$y = f(\succsim_1,\dots,\succsim_{h-1},\succsim'''_h,\succsim_{h+1},\dots,\succsim_I),$$

we have that $x \succsim''_h z$ implies $x \succsim'''_h z$ for every $z \in X$, and yet $y \neq x$.

There are two possibilities: Either $y \succ''_h x$ or $x \succsim''_h y$.

If $y \succ''_h x$ then the no-incentive-to-misrepresent condition is violated for the "true" preference relation $\succsim_h = \succsim''_h$ and the misrepresentation $\succsim'_h = \succsim'''_h$.

If $x \succsim''_h y$ then $x \succsim'''_h y$. Therefore, since no two distinct alternatives can be indifferent, $x \succ'''_h y$. But if $x \succ'''_h y$ then the no-incentive-to-misrepresent condition is violated for the "true" preference relation $\succsim_h = \succsim'''_h$ and the misrepresentation $\succsim'_h = \succsim''_h$. ■

REFERENCES

Arrow, K. J. (1963). *Social Choice and Individual Values*. 2d ed. New York: Wiley.

Austen-Smith, D., and J. S. Banks. (1996). *Positive Political Theory*. Ann Arbor: University of Michigan Press.

Caplin, A., and B. Nalebuff. (1988). On 64%-majority voting. *Econometrica* **56**: 787–814.

Grandmont, J-M. (1978). Intermediate preferences and majority rule. *Econometrica* **46**: 317–30.

May, K. (1952). A set of independent, necessary and sufficient conditions for simple majority decision. *Econometrica* **20**: 680–84.

Moulin, H. (1988) *Axioms of Cooperative Decision Making*. Cambridge, U.K.: Cambridge University Press.

Sen, A. (1970). *Individual Choice and Social Welfare*. San Francisco: Holden Day.

Sen, A. (1986). Social choice theory. Chap. 22 in *Handbook of Mathematical Economics*, edited by K. Arrow, and M. Intriligator. Amsterdam: North-Holland.

Shepsle, K. A., and M. Boncheck. (1995). *Analyzing Politics*. New York: W. W. Norton.

Tullock, G. (1967). The general irrelevancy of the general possibility theorem. *Quarterly Journal of Economics* **81**: 256–70.

EXERCISES

21.B.1[A] Verify that majority voting between two alternatives satisfies the properties of symmetry among agents, neutrality between alternatives, and positive responsiveness.

21.B.2[A] For each of the three properties characterizing majority voting between two alternatives according to Proposition 21.B.1 (symmetry among agents, neutrality between alternatives, and positive responsiveness) exhibit an example of a social welfare functional $F(\alpha_1, \ldots, \alpha_I)$ distinct from majority voting and satisfying the other two properties. This shows that none of the three properties is redundant for the characterization result.

21.B.3[A] Suppose there is a public good project that can take two levels $k \in \{0, 1\}$, where $k = 0$ can be interpreted as the status quo. The cost, in dollars, of any level of the public good is zero. There is a population I of agents having quasilinear preferences (with dollars as numeraire) over levels of the public good and money holdings. Thus, the preferences of agent i are completely described by the willingness to pay $v_i \in \mathbf{R}$ for the level $k = 1$ over the level $k = 0$. The number v_i may be negative (in this case it amounts to the minimum compensation required).

Show that a majority rule decision over the two levels of the public project guarantees a Pareto optimal decision over the set of policies constituted by the two levels of the public project (with no money transfers taking place across agents) but not over the larger set of policies in which transfers across agents are also possible.

Compare and contrast the majority decision rule (a "median") with the Pareto optimum decision rule for the case in which transfers across agents are possible (a "mean").

21.C.1[A] Provide the requested completion of step 1 of the proof of Proposition 21.C.1.

21.C.2[B] We can list the implicit and explicit assumptions of the Arrow impossibility theorem (Proposition 21.C.1) to be the following:

(a) The number of alternatives is at least 3.
(b) Universal domain: To be specific, the domain of the social welfare functional $F(\cdot)$ is $\mathscr{R}^I$.
(c) Social rationality: That is, $F(\succsim_1, \ldots, \succsim_I)$ is a rational preference relation (i.e. complete and transitive) for every possible profile of individual preferences.
(d) Pairwise independence (Definition 21.C.3).
(e) Paretian condition (Definition 21.C.2).
(f) No dictatorship: That is, there is no agent h that at any profile of individual preferences imposes his strict preference over any possible pair of alternatives (see Proposition 21.C.1 for a precise definition).

For each of these six assumptions exhibit a social welfare functional $F(\cdot)$ satisfying the other five. This shows that none of the conditions is redundant for the impossibility result.

21.C.3[A] Show that there are social welfare functionals $F: \mathscr{R}^I \to \mathscr{R}$ defined on $\mathscr{R}^I$ (i.e., individual indifference is possible) satisfying all the conditions of Arrow's impossibility theorem (Proposition 21.C.1) and for which, however, the social preferences are not *identical* to the preferences of any individual. [*Hint*: Try a *lexical dictatorship* in which the nth-ranked dictator imposes his preference if and only if every higher ranked dictator is indifferent.]

21.D.1[B] Suppose that X is a finite set of alternatives. Construct a reflexive and complete preference relation $\succsim$ on X with the property that $\succsim$ has a maximal element on every strict subset $X' \subset X$, and yet $\succsim$ is not acyclic.

21.D.2[A] Verify that the social preferences generated by the oligarchy example (Example 21.D.1) are quasitransitive but that social indifference may not be transitive. Interpret.

21.D.3[A] Show that the social preferences generated by the vetoers example (Example 21.D.2) are acyclic but not necessarily quasitransitive. Show also that in spite of the veto power of agent 2 it may happen that alternative x is the only maximal alternative for the social preferences.

21.D.4[A] With reference to Example 21.D.4, show that a continuous preference relation $\succsim$ on $X = [0, 1]$ is single peaked only if it is *strictly convex*.

21.D.5[A] Give a direct proof that none of the six linear orders possible among three alternatives can make the three preferences involved in the Condorcet paradox (Example 21.C.2) into a single-peaked family.

21.D.6[B] Give an example with an even number of agents and single-peaked preferences in which pairwise majority voting fails to generate a fully transitive social preference relation.

21.D.7[C] Suppose that X is a convex subset of $\mathbf{R}^2$ with the origin in its interior. There are three agents $i = 1, 2, 3$. Every i has a continuously differentiable utility function $u_i: X \to \mathbf{R}$. Assume that the cone in $\mathbf{R}^2$ spanned by the set of gradients at the origin $\{\nabla u_1(0), \nabla u_2(0), \nabla u_3(0)\}$ is the entire $\mathbf{R}^2$. Show the following:

(a) There are three alternatives $x, y, z \in X$ that constitute a Condorcet cycle (i.e., there is a strict majority for x over y, y over z, and z over x).
(b) (Harder) Given any $x \in \mathbf{R}^2$, there is a $y \in \mathbf{R}^2$ such that $\|x - y\| < \|x\|$ and y is preferred by two agents to the origin $0 \in \mathbf{R}^2$. That is, if you think of the origin as the status-quo then for any x we can find a strict majority that prefers, over the status-quo, an alternative that moves us closer to x. [*Hint:* You can safely assume that the utility functions are linear.]

21.D.8[C] The situation is as in Exercise 21.D.7 except that now, at the origin, the gradients of the utility functions constitute a pointed cone (i.e. the cone does not contain any half-space). Assume also that utility functions are quasiconcave.

(a) Argue that at the origin there is an agent who is a directional median in the sense that any alternative having a strict majority against the origin must make this agent strictly better off.
(b) Suppose now that at every $x \in X$ the cone spanned by $\{\nabla u_1(x), \nabla u_2(x), \nabla u_3(x)\}$ is pointed. Then according to (a) there is a directional median agent at every $x \in X$. Show that this directional median agent can change with x and that Condorcet cycles are possible.
(c) The situation is as in (b). Show that, if the directional median agent is the same at every $x \in X$, then there can be no Condorcet cycle.

21.D.9[C] (Grandmont) Consider a set of alternatives X. Given three rational preference relations $\succsim, \succsim', \succsim''$ on X, one says that $\succsim''$ is *intermediate* between $\succsim$ and $\succsim'$ if $x \succsim y$ and $x \succsim' y$ implies $x \succsim'' y$. That is, for every alternative y the intersection of the upper contour sets for $\succsim$ and $\succsim'$ is contained in the upper contour set for $\succsim''$.

(a) Show that if $u(x)$ and $u'(x)$ are utility functions for preferences on X then, for any positive numbers γ and ψ, the preference relation represented by $\psi u(x) + \gamma u'(x)$ is intermediate between the preference relations represented by $u(x)$ and $u'(x)$.
(b) Suppose we are given N functions $h_1(x), \ldots, h_N(x)$ defined on X. The preferences of agents are represented by utility functions of the form $u_\beta(x) = \beta_1 h_1(x) + \cdots + \beta_N h_N(x)$, where $\beta = (\beta_1, \ldots, \beta_N) \in \mathbf{R}^N_{++}$. Show that for any two alternatives $x, y \in X$, the set $B(x, y) = \{\beta \in \mathbf{R}^N_{++}: u_\beta(x) > u_\beta(y)\}$ is the intersection of $\mathbf{R}^N_{++}$ with a translated half-space.

(c) Argue that the conclusion from (b) is still correct if a parametrization of utility functions $u_\beta(x)$ by a $\beta \in \mathbf{R}^N$ is such that whenever β'' is a convex combination of β and β' then the preferences represented by $u_{\beta''}(x)$ are intermediate between the preferences represented by $u_\beta(x)$ and $u_{\beta'}(x)$.

(d) Continuing with the parametrization of (b), suppose that we take the limit situation where the population of agents is represented by a density $g(\beta)$ over $\mathbf{R}^N_{++}$. We say that β^* is a *median agent* for $g(\cdot)$ if *every* hyperplane in $\mathbf{R}^N$ passing through β^* divides $\mathbf{R}^N$ into two regions having equal mass according to the density $g(\cdot)$. Show that a median agent for an arbitrary $g(\cdot)$ may or may not exist.

(e) In the framework of (d), suppose there is a median agent β^*, that $g(\beta^*) > 0$, and that x^* is the single most preferred alternative of the median agent. Show then that x^* defeats any other alternative in pairwise majority voting.

(f) Show that the Euclidean preferences in Example 21.D.6 can be put into the framework of this exercise by keeping the sets of alternatives and of agents conceptually separated.

21.D.10[B] The purpose of this exercise is to illustrate the use of single-peakedness in a policy problem. Specifically, we consider the problem of determining by majority voting a tax level for wealth redistribution.

Suppose that there is an odd number I of agents. Each agent has a level of wealth $w_i > 0$ and an increasing utility function over wealth levels. The mean wealth is $\bar{w}$, and the median wealth is w^*.

(a) Interpret the distributional significance of a difference between $\bar{w}$ and w^*.

(b) Consider a proportional tax rate $t \in [0, 1]$ identical across agents. The set of alternatives is $X = [0, 1]$, the set of possible levels of the tax rate. Tax receipts are redistributed uniformly. Thus, for a tax rate t, the after-tax wealth of agent i is $(1 - t)w_i + t\bar{w}$. Show that the preferences over X of all agents are single peaked and that the Condorcet winner t_c is $t_c = 0$ or $t_c = 1$ according to whether $w^* > \bar{w}$ or $w^* < \bar{w}$, respectively. Interpret.

(c) Now suppose that taxation gives rise to a deadweight loss. Being very crude about it, suppose that a tax rate of $t \in [0, 1]$ decreases the pretax level of agent i's wealth to $w_i(t) = (1 - t)w_i$ [thus, the average tax receipts are $t(1 - t)\bar{w}$ and the ex post wealth level of agent i is $(1 - t)^2 w_i + t(1 - t)\bar{w}$]. Show that preferences on wealth levels are again single peaked (but notice that the after-tax wealth level may not be a concave function of the tax rate). Show then that $t_c \le \frac{1}{2}$. Also, $t_c = 0$ or $t_c > 0$ according to whether $w^* > \frac{1}{2}\bar{w}$ or $w^* < \frac{1}{2}\bar{w}$, respectively. Compare with (b) and interpret.

(d) Let us modify (c) by assuming that the deadweight loss affects individual wealth differently: A tax rate of $t \in [0, 1]$ decreases pretax wealth of agent i to $(1 - t^2)w_i$ [this is theoretically more satisfactory than the situation in (c) since we know from first principles that at $t = 0$ a small increase in t should have a second-order effect on total welfare]. Show then that individual preferences on tax rates need no longer be single peaked.

21.D.11[B] Consider a finite set of alternatives X and a set of preferences $\mathscr{P}_{\ge}$, single-peaked with respect to some linear order $\ge$ on X (note that we rule out the possibility of individual indifference). The number of agents is odd. As we have seen in Proposition 21.D.2, a possible class of social welfare functionals $F: \mathscr{P}^I_{\ge} \to \mathscr{P}$ that satisfy the Paretian and pairwise independence conditions are those where we fix a subset $S \subset I$ composed of an odd number of agents (a kind of oligarchy) and let the members of this subset determine social preferences by pairwise majority voting. Show by example that this is *not* the only possible class of social welfare functionals $F: \mathscr{P}^I_{\ge} \to \mathscr{P}$ that satisfy the Paretian and pairwise independence conditions.

21.D.12[A] Suppose that the total cost $c > 0$ of a project has to be financed by levying taxes from three agents. Therefore, the set of alternatives is $X = \{(t_1, t_2, t_3) \ge 0: t_1 + t_2 + t_3 = c\}$. The financing scheme is to be decided by majority voting.

(a) Show that no strictly positive alternative $(t_1, t_2, t_3) \gg 0$ can be a Condorcet winner.

(b) Discuss what happens with alternatives (t_1, t_2, t_3) where $t_i = 0$ for some i.

21.D.13[B] We have a population of agents (to be simple, a continuum) with Euclidean preferences in $\mathbf{R}^2$. The preferences of the agents fall into a finite number J of types. Each type is indexed by the most preferred point x_j. We assume that the x_j's are in "general position," in the sense that no three of the x_j's line up into a straight line. We denote by $\alpha_j \in [0, 1]$ the fraction of the total mass of agents that are of type j.

(a) Suppose that J is odd and $\alpha_1 = \cdots = \alpha_J$. Prove that if $y \in \mathbf{R}^2$ is a Condorcet winning alternative, then $y \in \{x_1, \ldots, x_J\}$. That is, the Condorcet winning alternative must coincide with the top alternative of some type. Does this remain true if J is even?

(b) (De Marzo) Suppose now that there is a dominant type, that is, a type h such that $\alpha_h > \alpha_j$ for every $j \ne h$. Prove that if there is a Condorcet winning alternative $y \in \mathbf{R}^2$, then $y = x_h$. That is, only the top alternative of the dominant type can be a Condorcet winning alternative.

21.D.14[B] In this exercise we verify that we cannot weaken the definition of single-peakedness to require only that preference be weakly increasing as the peak is approached.

Suppose we have five agents and five alternatives $\{x, y, z, v, w\}$. The individual preferences are

$$x \succ_1 y \sim_1 z \sim_1 v \sim_1 w,$$

$$y \succ_2 x \succ_2 z \succ_2 v \succ_2 w,$$

$$z \succ_3 y \sim_3 v \sim_3 w \succ_3 x,$$

$$v \succ_4 w \succ_4 x \sim_4 y \sim_4 z,$$

$$w \succ_5 x \sim_5 y \sim_5 z \sim_5 v,$$

(a) Show that there is no Condorcet winner among these alternatives; that is, every alternative is defeated by majority voting by some other alternative.

(b) Show that there is a linear order $\ge$ on the alternatives such that the preference relation of the five agents satisfies the following property: "Preferences are weakly increasing as we approach, in the linear order $\ge$, the most preferred alternative of the agent."

(c) Verify that the alternatives could be viewed as points in $[0, 1]$ and that the preferences of each agent could be induced by the restriction to the set of alternatives of a quasiconcave utility function on $[0, 1]$. [Note: $u_i(t)$ is quasiconcave if $\{t \in [0, 1]: u_i(t) \ge \gamma\}$ is convex for every γ.]

(d) (Harder) Extend the previous arguments and constructions into an example with the following characteristics: (i) There are five agents; (ii) the space of alternatives equals the interval $[0, 1]$; (iii) every agent has a quasiconcave utility function on $[0, 1]$ with a single maximal alternative; and (iv) there is no Condorcet winner in $[0, 1]$.

21.E.1[A] Consider a finite set of alternatives X and suppose that there is an odd number of agents. The domain of preferences is $\mathscr{A} = \mathscr{P}^I_{\ge}$, where $\ge$ is a linear order on X (i.e., preferences are single peaked and individual indifferences do not arise). Show that the social choice function that assigns the Condorcet winner to every profile satisfies the weak Pareto and the monotonicity conditions.