Chapter 12: Social Choice Theory

Felix Munoz-Garcia and Ana Espinola-Arredondo
School of Economic Sciences
Washington State University

1 Introduction

In this chapter, we consider a society with $N \geq 2$ individuals, each of them endowed with his own preference relation over a set of alternatives, such as the political candidates running for office, or the projects being considered for implementation in a region. We then examine the aggregation of individual preferences into a social preference, which can be interpreted as a ranking of alternatives for the society.

We discuss different methods of aggregating individual preferences and their properties, first in binary sets of alternatives, and then in sets with more than two alternatives (e.g., seventeen candidates competing in the 2015 primary of the Republican party to become the party's nominee for U.S. President in the general election). We explore the question, originally posed by Arrow (1953), of whether there exists a procedure to aggregate individual preferences satisfying a list of desirable, and not very demanding, properties (spoiler alert: it doesn't exist!).

We then analyze the reactions of the literature to Arrow's inexistence result, which can be divided into two groups: those studies that restricted the type of preferences that individuals can sustain; and those aggregating individual preferences into a social welfare function with a cardinal measure (rather than an ordinal ranking of social preferences). We finish this chapter discussing different voting procedures often used in real life, and their properties.

2 Social welfare functional

Consider a group of $N \geq 2$ individuals and a set of alternatives X. For simplicity, we first consider that set X is binary, thus containing only two elements $X = \{x, y\}$, and later on extend our analysis to sets with more than two alternatives. These two alternatives could represent, for instance, two candidates competing for office, or two policies to be implemented (i.e., the status quo and a new policy). In this binary setting, every individual i's preference over alternatives x and y can be defined as a number:

$$\alpha_i = \{1, 0, -1\}$$

indicating that he strictly prefers x to y when $\alpha_i = 1$, he is indifferent between x and y when $\alpha_i = 0$, or he strictly prefers alternative y to x when $\alpha_i = -1$. Our goal in this chapter is to aggregate individual preferences with the use of a social welfare functional (also referred to as social welfare aggregator); as defined below.

Social welfare functional. A social welfare functional (swf) is a rule

$$F(\alpha_1, \alpha_2, ..., \alpha_N) \in \{1, 0, -1\}$$

which, for every profile of individual preferences $(\alpha_1, \alpha_2, ..., \alpha_N) \in \{1, 0, -1\}^N$, assigns a social preference $F(\alpha_1, \alpha_2, ..., \alpha_N) \in \{1, 0, -1\}$.

As an example, for individual preferences $(\alpha_1, \alpha_2, \alpha_3) = (1, 0, 1)$ where individuals 1 and 3 strictly prefer x over y while individual 2 is indifferent, the swf could be $F(1, 0, 1) = 1$, thus preferring alternative x over y at the aggregate level. We next describe different properties of swfs, and test whether commonly used methods to aggregate individual preferences (that is, common swfs) satisfy or violate such properties.

2.1 Properties of swf

Paretian. For any pair of alternatives x and y in X, if $x \succ^i y$ for every individual i, then the social preference is $x \succ y$.

That is, when all individuals strictly prefer alternative x to y, we say that a swf is Paretian if it yields that x is strictly preferred by society, $F(1, 1, ..., 1) = 1$. Similarly, when all individuals strictly prefer alternative y to x, it yields that y is strictly preferred by society $F(-1, -1, ..., -1) = -1$. Note that we require that all individuals share the same strict preference for one of the two alternatives, which is then respected by the social ranking that the swf produces.[1]

This property is not very demanding. To see this point, note that a swf violating the Paretian property would pick alternative y as socially preferred even if all individuals strictly prefer alternative x to y! This property is, hence, satisfied by many swfs. For instance, simple and weighted voting, and even dictatorship, satisfy this property. We define these three swfs below and then confirm that they satisfy the Paretian property.

Example 12.1. Weighted voting swf. According to this swf, we first add individual preferences, assigning a weight $\beta_i \geq 0$ to every individual, where $(\beta_1, \beta_2, ..., \beta_N) \neq 0$, as in a weighted sum $\sum_i \beta_i \alpha_i \in \mathbb{R}$. Intuitively, the preference of every individual i, $\alpha_i = \{1, 0, -1\}$, is weighted by the importance that such society assigns to his preference, as captured by parameter β_i. After finding the weighted sum $\sum_i \beta_i \alpha_i$, we apply the sign operator, which yields 1 when the weighted sum is positive $\sum_i \beta_i \alpha_i > 0$, 0 when the weighted sum is zero $\sum_i \beta_i \alpha_i = 0$, and -1 when it is negative $\sum_i \beta_i \alpha_i < 0$. Hence, the weighted voting swf can be summarized as

$$F(\alpha_1, \alpha_2, ..., \alpha_N) = \text{sign} \sum_i \beta_i \alpha_i$$

[1] For this reason, the literature often refers to the property by saying that the swf respects "unanimity of strict preference."

In order to check if this swf is Paretian, we only need to confirm that, when all individuals strictly prefer alternative x to y, the swf also prefers x to y. Indeed,

$$F(1,1,...,1)=1, \text{ since } \sum_i \beta_i\alpha_i=\sum_i \beta_i>0;$$

and similarly, when all individuals strictly prefer alternative y to x, the swf also prefers y to x,

$$F(-1,-1,...,-1)=-1, \text{ since } \sum_i \beta_i\alpha_i=-\sum_i \beta_i<0. \quad \square$$

Example 12.2. Simple majority. This aggregation method is a special case of weighted majority, where the swf assigns the same weights to all individuals, i.e., $\beta_i=1$ for every individual i. In this setting, if the number of individuals who prefer alternative x to y is larger than the number of individuals preferring y to x, then the swf prefers x over y, i.e., $F(\alpha_1,\alpha_2,...,\alpha_N)=1$. The opposite argument applies if the number of individuals who prefer alternative y to x is larger than the number of individuals preferring x to y, where the swf aggregating individual preferences prefers y over x, i.e., $F(\alpha_1,\alpha_2,...,\alpha_N)=-1$. Finally, note that the simple majority swf is Paretian since weighted voting is Paretian too (as shown in Example 12.1). We nonetheless test this property as a practice:

$$\begin{aligned} F(1,1,...,1) &= 1, \text{ since } \sum_i \beta_i\alpha_i=N>0; \text{ and} \\ F(-1,-1,...,-1) &= -1, \text{ since } \sum_i \beta_i\alpha_i=-N<0. \quad \square \end{aligned}$$

Example 12.3. Dictatorial swf. We say that a swf is *dictatorial* if there exists an agent d, called the dictator, such that, for any profile of individual preferences $(\alpha_1,\alpha_2,...,\alpha_N)$:

1. when the dictator strictly prefers alternative x to y, $\alpha_d=1$, the swf also prefers x to y, $F(\alpha_1,\alpha_2,...,\alpha_N)=1$; and
2. when the dictator strictly prefers alternative y to x, $\alpha_d=-1$, the swf also prefers y to x, $F(\alpha_1,\alpha_2,...,\alpha_N)=-1$.

Intuitively, the strict preference of the dictator α_d prevails as the social preference, regardless of the preference profile of all other individuals $i\neq d$. We can, hence, understand the dictatorial swf as a extreme case of weighted voting where the weight assigned to individual d's preferences is positive, $\beta_d>0$, but nil for all other individuals in the society, i.e., $\beta_i=0$ for all $i\neq d$. Finally, note that since the weighted voting swf is Paretian, then the dictatorial swf (as a special case of weighted voting) must also be Paretian. For completeness, we show it next:

$$\begin{aligned} F(1,1,...,1) &= 1, \text{ since } \sum_i \beta_i \alpha_i = \beta_d > 0; \text{ and} \\ F(-1,-1,...,-1) &= -1, \text{ since } \sum_i \beta_i \alpha_i = -\beta_d < 0. \quad \square \end{aligned}$$

After confirming that weighted majority (and its two special cases, simple majority and dictatorship) satisfy the Paretian property, we continue our presentation of properties that are commonly regarded as desirable for voting procedures to satisfy.

Symmetry among voters (or anonymity). The swf $F(\alpha_1, \alpha_2, ..., \alpha_N)$ 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. More precisely, let

$$\pi : \{1, 2, ..., N\} \to \{1, 2, ..., N\}$$

be an onto function (i.e., a function that, for every individual i, identifies another individual j such that $\pi(j) = i$). Then, for every profile of individual preferences $(\alpha_1, \alpha_2, ..., \alpha_N)$, the swf prefers the same alternative with $(\alpha_1, \alpha_2, ..., \alpha_N)$ and with the "permuted" profile of individual preferences $\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}, ..., \alpha_{\pi(N)}\right)$, that is,

$$F(\alpha_1, \alpha_2, ..., \alpha_N) = F\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}, ..., \alpha_{\pi(N)}\right)$$

Intuitively, the name of the individual preferring x over y, preferring y over x, or being indifferent between them, does not affect the socially preferred alternative. Informally, this property says that the candidate chosen by a swf should depend only on voters' preferences and not *who* has those preferences. That is, voters should be treated symmetrically when aggregating their preferences. As a practice, Example 12.4 below shows that simple majority satisfies anonymity, but weighted voting or dictatorship do not.

Example 12.4. Testing for anonymity. *Simple majority.* Let us first check if simple majority satisfies anonymity. The sum $\sum_i \alpha_i$ coincides when we consider the initial preference profile $(\alpha_1, \alpha_2, ..., \alpha_N)$ and when we consider the "permuted" profile of individual preferences $\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}, ..., \alpha_{\pi(N)}\right)$; thus yielding the same social preference when society uses simple majority to aggregate individual preferences.[2]

Weighted majority. The sum $\sum_i \beta_i \alpha_i$ that we obtain when aggregating individual preferences according to weighted voting, however, can differ in the initial and permuted preference profile.

[2] For instance, consider a profile of individual preferences $(\alpha_1, \alpha_2) = (1, 0)$ and its permutation $\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}\right) = (0, 1)$ where $\pi(1) = 2$ and $\pi(2) = 1$, i.e., individual preferences are switched between individuals 1 and 2. Then, the simple majority swf yields x as being socially preferred according to both the initial preference profile $(\alpha_1, \alpha_2) = (1, 0)$, $F(1, 0) = 1$ since $\sum_i \alpha_i = 1 + 0 = 1$, and according to the permuted preference profile $\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}\right) = (0, 1)$, i.e., $F(0, 1) = 1$ since $\sum_i \alpha_i = 0 + 1 = 1$.

Consider, for instance, a society of three individuals with a profile of individual preferences given by $(\alpha_1, \alpha_2, \alpha_3) = (1, -1, 1)$. According to weighted voting, we find that

$$\sum_i \beta_i \alpha_i = \beta_1 - \beta_2 + \beta_3,$$

which is positive if $\beta_1 > \beta_2 - \beta_3$. In contrast, if we apply the following permutation to individuals' idientities, $\pi(1) = 3$, $\pi(2) = 1$, and $\pi(3) = 2$, weighted voting now yields

$$\sum_i \beta_i \alpha_i = -\beta_1 + \beta_2 + \beta_3,$$

which is positive if $\beta_1 < \beta_2 + \beta_3$. Therefore, weighted voting yields the same output before and after the permutation (thus satisfying anonymity) only if β_1 is intermediate, i.e., $\beta_2 - \beta_3 < \beta_1 < \beta_2 + \beta_3$; but it yields different outcomes otherwise (thus violating anonymity).

Dictatorship. A similar argument applies to the dictatorial swf since it is a special case of weighted voting where $\beta_d > 0$ for the dictator and $\beta_i = 0$ for all other individuals $i \neq d$. For instance, if in the above society with three individuals, $\beta_1 = 1$ and $\beta_2 = \beta_3 = 0$, then the initial profile of individual preferences $(\alpha_1, \alpha_2, \alpha_3) = (1, -1, 1)$ yields $\sum_i \beta_i \alpha_i = 1$ (which entails alternative x as strictly preferred to y by the society), whereas the "permuted" profile of individual preferences $\left(\alpha_{\pi(1)}, \alpha_{\pi(2)}, \alpha_{\pi(3)}\right) = (-1, 1, 1)$ yields $\sum_i \beta_i \alpha_i = -1$ (meaning that society now strictly prefers alternative y to x). This should come at no surprise since individual 1 is the dictator, thus imposing his strict preference for x in the initial profile of preferences. However, individual 2 becomes individual 1 (and thus the dictator) after we permute the identities of these individuals, allowing him to impose his strict preference of y over x. □

Neutrality between alternatives. The swf $F(\alpha_1, \alpha_2, ..., \alpha_N)$ is *neutral between alternatives* if, for every profile of individual preferences $(\alpha_1, \alpha_2, ..., \alpha_N)$,

$$F(\alpha_1, \alpha_2, ..., \alpha_N) = -F(-\alpha_1, -\alpha_2, ..., -\alpha_N)$$

That is, if we reverse the preferences of all agents, from $(\alpha_1, \alpha_2, ..., \alpha_N)$ to $(-\alpha_1, -\alpha_2, ..., -\alpha_N)$, then the social preference is reversed as well. For instance, if the profile of individual preferences is $(\alpha_1, \alpha_2) = (1, 0)$ and the swf produces $F(1, 0) = 1$, then when we reverse the profile of individual preferences to $(-1, 0)$, the social preference must become $F(-1, 0) = -1$ for the swf to satisfy neutrality. Intuitively, this property is often understood as that the swf treats alternatives x and y symmetrically, without providing an initial advantage to either alternative. As a practice, check that simple majority voting satisfies neutrality between alternatives.

Positive responsiveness. Consider a profile of individual preferences $(\alpha_1, \alpha_2, ..., \alpha_N)$ where alternative x is socially preferred or indifferent to y, i.e., $F(\alpha_1, \alpha_2, ..., \alpha_N) \geq 0$. Take now a new profile $(\alpha'_1, \alpha'_2, ..., \alpha'_N)$ in which some agents raise their consideration for x, i.e., $(\alpha'_1, \alpha'_2, ..., \alpha'_N) \geq$

$(\alpha_1, \alpha_2, ..., \alpha_N)$ where $(\alpha'_1, \alpha'_2, ..., \alpha'_N) \neq (\alpha_1, \alpha_2, ..., \alpha_N)$, i.e., $\alpha'_i > \alpha_i$ for at least one individual i. We say that a swf is *positively responsive* if the new profile of individual preferences $(\alpha'_1, \alpha'_2, ..., \alpha'_N)$ makes alternative x socially preferred, i.e., $F(\alpha'_1, \alpha'_2, ..., \alpha'_N) = 1$.

In words, initial condition $(\alpha'_1, \alpha'_2, ..., \alpha'_N) \geq (\alpha_1, \alpha_2, ..., \alpha_N)$ says that at least one of the components of the new preference profile is larger than in the original profile, thus indicating that the consideration of alternative x increased for at least one individual. Therefore, positive responsiveness says two things:

1. If alternative x was socially preferred to y under the initial profile of preferences,

$$F(\alpha_1, \alpha_2, ..., \alpha_N) = 1,$$

then x must still be socially preferred under the new profile of preferences, $F(\alpha'_1, \alpha'_2, ..., \alpha'_N) = 1$; and

2. If alternative x was indifferent to y under the initial profile of preferences,

$$F(\alpha_1, \alpha_2, ..., \alpha_N) = 0,$$

then x becomes socially preferred under the new profile of individual preferences, $F(\alpha'_1, \alpha'_2, ..., \alpha'_N) = 1$.

As you probably suspect, point (1) is not demanding, but point (2) does not hold for some voting procedures. To see this point, consider an extreme case where every individual is indifferent between x and y in the initial profile of preferences, yielding $F(\alpha_1, \alpha_2, ..., \alpha_N) = 0$. According to positive responsiveness, if at least one individual strictly prefers x to y in the new preference profile, while everyone else is still indifferent between both alternatives, society should strictly prefer x, that is, $F(\alpha'_1, \alpha'_2, ..., \alpha'_N) = 1$. Example 12.5 presents a swf satisfying this property while Example 12.6 tests this property in common swfs.

Example 12.5. A swf satisfying positive responsiveness. Consider an individual preference profile of $(\alpha_1, \alpha_2, \alpha_3) = (1, 0, -1)$, and assume that the swf in this case yields $F(\alpha_1, \alpha_2, \alpha_3) = 0$, i.e., society is indifferent between alternatives x and y. Now, assume that the preference profile increases the consideration of alternative x, implying that the new preference profile can be any of the following[3]

$$\begin{aligned}(\alpha_1', \alpha_2', \alpha_3') &= (1, 1, -1) \\ &= (1, 0, 0) \\ &= (1, 0, 1) \\ &= (1, 1, 0) \\ &= (1, 1, 1)\end{aligned}$$

Then, if the swf selects $F(\alpha_1', \alpha_2', \alpha_3') = 1$, meaning that alternative x is socially preferred to y, the swf satisfies positive responsiveness. If, instead, the swf for the new preference profile is still $F(\alpha_1', \alpha_2', \alpha_3') = 0$ or decreases to $F(\alpha_1', \alpha_2', \alpha_3') = -1$, the such a swf violates this property. □

Example 12.6. Testing positive responsiveness. In this example, we show that weighted voting satisfies positive responsiveness, implying that simple majority and dictatorship must also satisfy it since they are both special cases of weighted voting. To test this property, we need to check the two conditions listed above.

1. First, when society is indifferent between alternative x and y under the initial preference profile, the weighted sum is $\sum_i \beta_i \alpha_i = 0$. In this context, the sum in the new preference profile, $\sum_i \beta_i \alpha_i'$, must be positive since at least one of the elements in the vector of individual preferences (that is, at least of the α_i's) increased. Therefore, the sum $\sum_i \alpha_i' > 0$ implying that society strictly prefers alternative x to y.
2. Second, when society strictly prefers alternative x to y under the initial preference profile, the weighted sum is $\sum_i \beta_i \alpha_i > 0$. In this case, the sum in the new preference profile, $\sum_i \beta_i \alpha_i'$, must also be positive, thus implying that x is socially preferred to y under the new profile of individual preferences as well.

Since conditions (1) and (2) hold, weighted voting satisfies positive responsiveness; and so do simple majority and dictatorship since they are special cases of weighted voting as discussed above. □

2.2 Arrow's impossibility theorem

Let us now extend our analysis to non-binary sets of alternatives X, e.g., three candidates competing for elected office or, more generally, $X = \{a, b, c, ...\}$. In this context, the aggregation of individual preferences using a majority voting swf, or a weighted voting swf, can be subject to non-transitivities

[3] In the first preference profile, individual 2 is the only agent increasing his consideration for alternative x (as $\alpha_2 = 0$ increases to $\alpha_2' = 1$); and in the second perference profile, individual 3 is the only agent increasing his consideration for x (since $\alpha_3 = -1$ increases to $\alpha_3' = 0$). In the third preference profile, individual 3 is the only agent whose consideration for x increases, but now α_3 increases from $\alpha_3 - 1$ to $\alpha_3' = 1$; while in the fourth and fifth preference profiles, the consideration of individuals 2 and 3 simultaneously increase.

in the resulting social preference. In other words, the order in which pairs of alternatives are voted can lead to cyclicalities, as shown in Condorcet's paradox (see Chapter 1, where we aggregated different criteria of a prospective student admitted into three Ph.D. programs). For illustration purposes, we recall Condorcet paradox below.

Condorcet Paradox. Consider a society with three individuals and three alternatives $X = \{x, y, z\}$, where individual preferences are given by

$$\begin{aligned} x &\succsim^1 y \succsim^1 z \quad \text{for individual 1,} \\ y &\succsim^2 z \succsim^2 x \quad \text{for individual 2, and} \\ z &\succsim^3 x \succsim^3 y \quad \text{for individual 3} \end{aligned}$$

Cyclicalities. If individual preferences are aggregated according to majority voting, the resulting social preference is intransitive. Indeed, the social preference is

$$x \succsim^{1,3} y \succsim^{1,2} z \succsim^{2,3} x$$

where the superscripts above each preference symbol represent the individuals who sustain these preferences. For instance, $x \succsim^{1,3} y$ indicates that individuals 1 and 3 prefer alternative x to y (with only individual 2 prefering y to x), and thus alternative x beats y under majority voting. A similar argument applies to $y \succsim^{1,2} z$, where alternative y beats z under majority voting since individuals 1 and 2 weakly prefer y to z; and to $z \succsim^{2,3} x$, where alternative z beats x given that individuals 2 and 3 weakly prefer it. After comparing every pair of alternatives according to majority voting, we obtain an intransitive social preference relation leading to cyclicalities.

Agenda manipulation. Cyclicalities are important because they produce social preferences that are subject to agenda manipulation. In words, this means that the individual selecting which two alternatives are considered first in a pairwise vote can strategically choose two alternatives to produce his most preferred alternative as the winner of the pairwise vote, i.e., she can alter the agenda on his own benefit. In order to see this result, note that if alternatives x and y are confronted using majority voting, alternative x wins as it receives two votes (from voters 1 and 3) while alternative y only receives voter 2's ballot. The winner of this pairwise majority voting, x, is then paired against the remaining alternative, z, which yields z as the winner, since z receives two votes (from voters 2 and 3) while alternative x only receives voter 1's ballot. Hence, if the parwise vote is first between alternatives x and y with the winner subsequently confronting alternative z, a sophisticated agenda setter could anticipate that alternative z will be declared the winner. (Individual 3, for instance, would have incentives to set such an agenda for pairwise votes.)[4]

[4] If, instead, alternatives x and z are paired first, with the winner confronting alternative y, the outcome changes. Indeed, z wins a pairwise vote against x (as it receives votes from 2 and 3), but the winner, z, then loses against the remaining alternative y (as y is preferred to z by voters 2 and 1), thus declaring y as the winner. This agenda would be beneficial for individual 2. A similar argument applies if alternatives y and z are first matched in a pairwise voting, and the winner of this pair confronting afterwards alternative x. In particular, y would beat z since voters 1

Given the possibility of intransitive social preferences when using common voting methods, such as simple and weighted majority, an interesting question is: Can we design voting systems (i.e., swfs that aggregate individual preferences) that are not prone to the Condorcet's paradox and satisfy a minimal set of "desirable" properties? This was the question Arrow asked himself (for his Ph.D. thesis) obtaining a rather grim result: such a voting procedure does not exist!. This result is commonly known as "Arrow's impossibility theorem." Before presenting the theorem, let us first define the four minimal requirements that Arrow considered all swfs should satisfy.

1. **Unrestricted domain (U).** The domain of the swf, $\left(\succsim^1, \succsim^2, ..., \succsim^N\right)$, must allow for all combinations of individual preference relations on X.

   In other words, we allow any sort of individual preferences over alternatives.

2. **Paretian (P).** For any pair of alternatives x and y in X, if $x \succ^i y$ for every individual i, then the social preference is $x \succ y$.

   This is the same property we defined in elections between two alternatives, extended now to settings with more than two candidates. In words, if every member of society strictly prefers alternative x to y, society strictly prefers x to y.[5]

3. **Non-dictatorship (ND).** There is no individual d such that, for every pair of alternatives $(x, y) \in X$, individual d's strict preference of x over y, $x \succ^d y$, implies a social preference of $x \succ y$ regardless of the preferences of all other individuals $i \neq d$.

   Note that this is a very mild assumption. Indeed, we could consider a "virtual" dictatorship in which an individual d imposes his preference on the rest of individuals for all, but one, pair of alternatives. Such a setting would be considered non-dictatorial, since for a dictatorship to arise we must find that the preferences of individual d dictate the social preference for *all* pairs of alternatives. Consider, for instance, a group of two individuals, with the profile of individual preferences depicted in table 12.1. While social preferences coincide with the preferences of individual 1 in most pairs of alternatives (see first two rows of the table), they do not in one pair (see last row). As a consequence, this swf is not dictatorial.

$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$
$x \succsim^1 y$	$y \succsim^2 x$	$x \succsim y$
$y \succsim^1 z$	$z \succsim^2 y$	$y \succsim z$
$z \succsim^1 w$	$w \succsim^2 z$	$w \succsim z$

Table 12.1. A "virtual" dictatorship of individual 1.

and 2 would vote for y, but the pairwise voting between y and the remaining alternative x ultimately yields x as the winner (since voters 1 and 3 vote prefer x over y). Such voting agenda would be particularly attractive to individual 1.

[5] Recall that, if all but one individual strictly prefers alternative x to y, yet one person is indifferent between these two alternatives, a Paretian swf does not necessarily yield that alternative x is strictly preferred to y.

4. **Independence of irrelevant alternatives (IIA).** Let $\succsim$ be social preferences arising from the list of individual preferences $\left(\succsim^1, \succsim^2, \dots \succsim^N\right)$, and $\succsim'$ that arising when individual preferences are $\left(\succsim'^1, \succsim'^2, \dots \succsim'^N\right)$. In addition, let x and y be any two alternatives in X. If each individual ranks alternatives x and y in the same way under his initial preferences, $\succsim^i$, and new preferences, $\succsim'^i$, then the social ranking of alternatives x and y should not change.

   Note that the premise of IIA only requires that, if individual i's preferences are $x \succsim^i y$ between alternatives x and y (e.g., his preference in the morning), then he keeps ranking these two alternatives in the same way, $x \succsim'^i y$ (e.g., his preferences in the afternoon), even if his preferences for a third alternative change.[6] Then, we say that a swf satisfies IIA if the social ranking of alternatives x versus y is unaffected. In other words, even if other alternatives different from x and y change their ranking when we move from $\succsim^i$ to $\succsim'^i$, individual preferences for x and y have not changed and, hence, the social preference for x and y should not change either. The following examples illustrates swfs satisfying or violating IIA.

Example 12.7. Swfs satisfying/violating IIA. Consider the preference profile depicted in table 12.2. Suppose that in the morning (left side) some individuals prefer alternative x to y, $x \succsim^i y$, while others prefer y to x, $y \succsim^i x$. However, they all rank alternative z below both x and y. In addition, suppose that the swf yields a social preference of x over y, i.e., $x \succsim y$. During the afternoon (right side), alternative z is ranked above both x and y for all individuals. However, the ranking of alternatives x and y did not change for any individual, i.e., if $x \succsim^i y$ then $x \succsim'^i y$, and if $y \succsim^i x$ then $y \succsim'^i x$; as required by the premise of IIA. Then, IIA says that society should still prefer x over y in the afternoon, i.e., $x \succsim' y$.

Morning			Afternoon		
$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$	$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$
.	.	.	z	.	z
.	.	.	.	z	.
x	y	x	x	y	x
y	x	y	y	x	y
.	z	z	.	.	.
z	.	.	.	.	.
$x \succsim^1 y$	$y \succsim^2 x$	$x \succsim y$	$x \succsim^1 y$	$y \succsim^2 x$	$x \succsim y$

Table 12.2. Swf satisfying IIA.

In contrast, table 12.3 illustrates a swf violating IIA. While the preferences over alternatives x and y of individuals 1 and 2 remain constant over time (i.e., the premise of IIA holds), the social

[6] The preferences of individual $j \neq i$ could, however, be different from those of individual i, i.e., $y \succsim^j x$ and $y \succsim'^j x$ so individual j ranks alternatives x and y in the same way in the morning and afternoon.

preference over these alternatives changes from $x \succsim y$ in the morning to $y \succsim x$ in the afternoon. Hence, IIA does not hold.

Morning				Afternoon		
$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$		$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$
.	.	.		z	.	z
.	.	.		.	z	.
x	y	x		x	y	y
y	x	y		y	x	x
.	z	z		.	.	.
z	.	.		.	.	.
$x \succsim^1 y$	$y \succsim^2 x$	$x \succsim y$		$x \succsim^1 y$	$y \succsim^2 x$	$y \succsim x$

Table 12.3. Swf violating IIA.

Note that if the preference between alternatives x and y changes for at least one individual from the morning to the afternoon, then the premise of IIA does not hold; as illustrated in table 12.4. In such a case, we cannot claim that IIA is violated.[7] □

Morning				Afternoon		
$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$		$\succsim^1$	$\succsim^2$	Soc.Pref. $\succsim$
.	.	.		z	.	z
.	.	.		.	z	.
x	y	x		x	x	y
y	x	y		y	y	x
.	z	z		.	.	.
z	.	.		.	.	.
$x \succsim^1 y$	$y \succsim^2 x$	$x \succsim y$		$x \succsim^1 y$	$x \succsim^2 y$	$y \succsim x$

Table 12.4. Swf for which the premise of IIA does not hold.

Most of these assumptions are often accepted as the minimal assumptions that we should impose on any swf that aggregates individual preferences into a social preference. Let us now describe Arrow's impossibility theorem, which comes as a surprising, even disturbing, result.

Arrow's impossibility theorem. *If there are at least three elements in the set of alternatives X, then there is no swf that simultaneously satisfies properties U, P, IIA, and ND.*

Proof: We will assume that U, P and IIA hold, and show that all swfs simultaneously satisfying these three properties must be dictatorial, thus violating one of the four properties. (U is used

[7] For us to claim that the swf violates the IIA, we first need that its premise to be satisfied, and the conclusion is violated, i.e., while individual preferences for x and y do not change throughout the day, the social preference for alternatives x and y changes.

throughout the proof when we alter the profile of individual preferences, since this property allows for all preference profiles to be admissible.)

Step 1: Consider that an alternative c is placed at the bottom of the ranking of every individual i. Then, by P, alternative c must be placed at the bottom of the social ranking as well; as depicted in the last column of table 12.5.

$\succsim^1$	$\succsim^2$	...	$\succsim^N$	Soc.Pref. $\succsim$
x	x'	...	x''	x'''
y	y'	...	y''	y'''
.	.			.
.	.			.
.	.			.
c	c	...	c	c

Table 12.5. Alternative c is at the bottoom of everyone's ranking.

Step 2: Imagine now that we move alternative c from the bottom of individual 1's ranking to the top of his ranking, leaving the position of all other alternatives unaffected. Next, we do the same move for individual 2, then for individual 3, etc.; as illustrated in table 12.6. Let individual n be the first for which raising alternative c to the top of his ranking causes the social ranking of alternative c to increase. The following table places alternative c at the *top* of the social ranking, a result we show next.

$\succsim^1$	$\succsim^2$	...	$\succsim^n$	...	$\succsim^N$	Soc.Pref. $\succsim$
c	c	...	c	...	x''	c
x	x'	...		...	y''	.
y	y'					.
.	.					.
.	.					.
.	.					.
w	w'	...		...	c	w'''

Table 12.6. Alternative c is raised to the top of the ranking for $i = 1, 2, ..., n$.

By contradiction, assume that the social ranking of c increases but *not to the top*, i.e., there is at least one alternative α such that $\alpha \succsim c$ and one alternative β for which $c \succsim \beta$, where $\alpha, \beta \neq c$; as depicted in table 12.7 (left panel). Because alternative c is either at the top of the ranking of individuals $1, 2, ..., n$, or at the bottom of the ranking of individuals $n + 1, ..., N$, we can change each individual i's preferences so that $\beta \succ^i \alpha$, while leaving the position of c unchanged for that individual; as depicted in the right panel of table 12.7.

$\succsim^1$	$\succsim^2$	...	$\succsim^n$	...	$\succsim^N$	Soc.Pref. $\succsim$
c	c	...	c	...	x''	α
x	x'	...		...	y''	.
y	y'					c
.	.					β
.	.					.
.	.					.
w	w'	...		...	c	w'''

$\succsim^1$	$\succsim^2$	...	$\succsim^n$	...	$\succsim^N$	Soc.Pref. $\succsim$
c	c	...	c	...	β	α
β	β	...	β	...	z	.
.	α					c
.	.				α	β
α	.					.
.	.		α			.
w	w'	...		...	c	w'''

Table 12.7. Alternative c must be at the top of the social ranking.

We have now changed each individual i's preferences so that $\beta \succ^i \alpha$, while leaving the position of c unchanged for that individual, which produces our desired contradiction:

1. On one hand, $\beta \succ^i \alpha$ for every individual which, by the P property, must yield a social preference of $\beta \succ \alpha$; and

2. On the other hand, the ranking of alternative α relative to c, and of β relative to c, have not changed for any individual.[8] By the IIA, this result implies that the social ranking of α relative to c, and of β relative to c, must remain unchanged. Hence, the social ranking is still $\alpha \succsim c$ and $c \succsim \beta$. By transitivity, this yields that $\alpha \succsim \beta$.

However, the result from point 1 ($\beta \succ \alpha$) contradicts that from point 2 ($\alpha \succsim \beta$); yielding the desired contradiction. Hence, alternative c must have moved all the way to the top of the social ranking. (Q.E.D.)

In the next, and final, step of the proof, we show that individual n is a dictator, thus imposing his preferences on the group regardless of the preference profile of all other individuals.

Step 3: Consider now two distinct alternatives a and b, each different from c. In Table 12.8, let's change the preferences of individual n as follows:

$$a \succ^n c \succ^n b$$

For every other individual $i \neq n$, we rank alternatives a and b in any way but keeping the position of c unchanged (see Table 12.8).

[8] That is, if an individual i's preference for alternative α and c was $\alpha \succsim^i c$ ($c \succsim^i \alpha$) before changing $\beta \succ^i \alpha$ for all individuals, his preference is still $\alpha \succsim^i c$ ($c \succsim^i \alpha$, respectively) after the change. A similar argument applies for individual i's preference between β and c, which remains unaffected after imposing the condition of $\beta \succ^i \alpha$ in the ranking of all individuals.

$\succsim^1$	$\succsim^2$	...	$\succsim^n$	...	$\succsim^I$	Soc.Pref. $\succsim$
c	c	...	a	...	x''	c
x	x'	...	c	...	y''	.
y	y'		b		.	.
.	.				.	a
.	.				.	b
a	b	...		...	a	.
b	a	...		...	b	
.	.					.
w	w'	...		...	c	w'''

Table 12.8. Individual n must be a dictator.

In the new profile of individual preferences, the ranking of alternatives a and c is the same for every individual as it was just *before* raising alternative c to the top of individual n's ranking in Step 2. Therefore, by IIA, the social ranking of alternatives a and c must be the same as it was at that moment (just *before* raising c to the top of individual n's ranking in Step 2). That is, $a \succ c$, since at that moment alternative c was at the bottom of the social ranking. Similarly, in the new profile of individual preferences, the ranking of alternatives c and b is the same for every individual as it was just after raising c to the top of individual n's ranking in Step 2. Hence, by IIA, the social ranking of alternatives c and b must be the same as it was at that moment. That is, $c \succ b$, since at that moment alternative c had just risen to the top of the social ranking. Summarizing, since $a \succ c$ and $c \succ b$, we have that, by transitivity, $a \succ b$.

Then, no matter how individuals different from individual n rank every pair of alternatives a and b, the social ranking agrees with individual n's ranking; thus showing that individual n is a dictator, which completes the proof.[9] (Q.E.D.)

In summary, we started with a swf satisfying properties U, P, and IIA, and showed that the social preference must coincide with that of one individual, thus violating the non-dictatorship property (ND). Other proofs of this theorem follow a similar route, by considering that three of the four assumptions hold, and showing that the fourth assumption must be violated.[10]

3 Reactions to Arrow's impossibility theorem

After Arrow's negative result to the search of a swf satisfying his four minimal assumptions, the literature reacted using two main approaches: (1) Eliminating the U assumption, by focusing on

[9] That is, while $a \succ^i b$ for some individuals and $b \succ^j a$ for other individuals, the fact that $a \succ^n b$ for individual n implies that $a \succ b$ for the social ranking, which is true for any two alternatives $a, b \neq c$, and regardless of how many individuals strictly prefer alternative a to b, and how many strictly prefer b to a.

[10] Other approaches use figures to provide a more visual representation of the proof; see, for instance, section 2.4 in Gaertner's (2009) book. Maskin and Sen (2014) also provide a combination of technical and intuition discussions on Arrow's impossibility theorem.

specific types of individual preferences, such as the single-peaked preferences that we define below; and (2) Aggregating the intensity of individual preferences (not only the ranking of alternatives for each individual) into a social welfare function. This last approach differs from the swf analyzed in previous sections, as that provided us with a ranking of social preferences (i.e., a cardinal measure), while a social welfare function evaluates the welfare that society achieves from each allocation of goods and services (i.e., an ordinal measure). We explore each approach in the next two subsections.

3.1 First reaction - Single-peaked preferences

We informally say that preferences of individual i are single-peaked if we can identify a blissing point (or satisfaction point) at which the individual reaches his maximal utility (i.e., his utility "peak"). Formally, the single peak is defined relative to a linear order $\geq$ on the set of available alternatives X.[11]

Single-peaked preferences *Definition***.** 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 the set of alternatives below x, $\{y \in X : x \geq y\}$, and decreasing with respect to $\geq$ on the set of alternatives above x, $\{y \in X : y \geq x\}$.

As suggested above, this definition can be understood as that there is an alternative x that represents a "peak" of satisfaction; and that satisfaction increases as we approach this peak either from points below x, $x \geq y$, or from points above x, $y \geq x$, so there cannot be other peak of satisfaction.

Example 12.8. Single-peaked preferences. Consider a set of policy alternatives $X = [0, 1]$, e.g., percentage of the federal budget spent in education. Every individual i's utility from alternative $x_k \in [0, 1]$ is

$$u(x_k, \theta_i) = -(x_k - \theta_i)^2$$

where $\theta_i \in [0, 1]$ represents individual i's ideal policy. To understand this utility function, note that it collapses to zero when the policy alternative coincides with the individual's policy ideal, $x_k = \theta_i$; but becomes a negative number both when the policy falls below his policy ideal, $x_k < \theta_i$, and when it exceeds his policy ideal, $x_k > \theta_i$. Graphically, the utility function exhibits an inverted U-shape, and lies in the negative quadrant for all $x_k \neq \theta_i$, and becomes zero only at $x_k = \theta_i$. This function is then single-peaked at $x_k = \theta_i$. $\square$

Example 12.9. Single-peaked preferences and convexity. Consider a set of alternatives $X = [a, b] \subset \mathbb{R}$, i.e., a segment of the real line. Then, a preference relation $\succsim$ on X is single peaked if

[11] Recall the definition of a linear order. We say that a binary relation $\geq$ is a *linear order* on the set of alternatives X if it is: (1) reflexive, i.e., $x \geq x$ for every $x \in X$; (2) transitive, i.e., $x \geq y$ and $y \geq z$ implies $y \geq z$; and (3) total, i.e., for any two distinct $x, y \in X$, we have that either $x \geq y$ or $y \geq x$, but not both. If the set of alternatives is a segment of the real line, i.e., $X \subset \mathbb{R}$, then the linear order $\geq$ can be understood as the "greater than or equal to" operator in the real numbers.

and only if it is *strictly convex*: That is if, for every alternative $w \in X$, and for any two alternatives y and z weakly preferred to w, i.e., $y \succsim w$ and $z \succsim w$ where $y \neq z$, their linear combination is strictly preferred to w,

$$\alpha y + (1-\alpha)z \succ w \quad \text{for all } \alpha \in (0,1)$$

For illustration purposes, Figure 12.1 depicts utility functions satisfying (violating) the single-peaked property in the left panel (right panel, respectively). In particular, note that in both panels $u(y) \geq u(w)$ and $u(z) \geq u(w)$; as required by the premise of convexity. The linear combination of y and z yields a utility $u(\alpha y + (1-\alpha)z)$ that lies above (below) $u(w)$ when the utility function has a single peak (multiple peaks); as depicted in the left (right) panel.

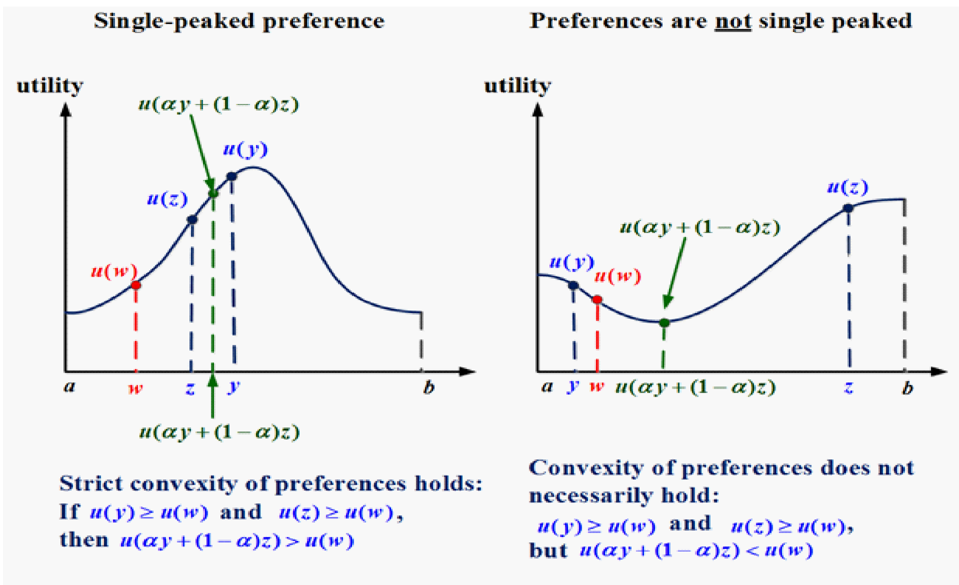

Figure 12.1. Single-peaked preferences and convexity.

Importantly, the single-peaked property is not equivalent to strict concavity in the utility function.[12] Figure 12.2 depicts a utility function that, despite being strictly convex, satisfies the single-peaked property. Indeed, $u(y) \geq u(w)$ and $u(z) \geq u(w)$, and the linear combination of y and z yields a utility $u(\alpha y + (1-\alpha)z)$ that lies above $u(w)$; as requred by the single-peaked property. $\square$

[12] This is analogous to say that strictly convex preferences are not equivalent to strictly concave utility functions. For more details on these properties, see Section 1.7 in Chapter 1.

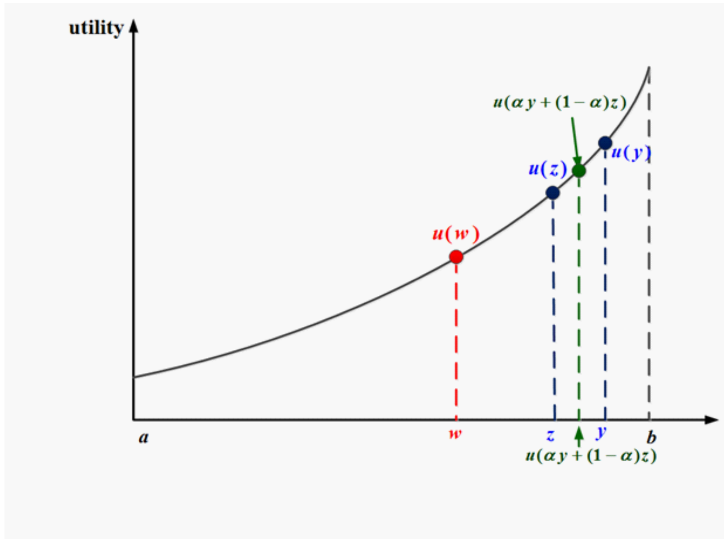

Figure 12.2. Single-peaked preferences and strictly concave utility function.

We will now restrict our attention to settings in which all individuals have single-peaked preferences with respect to the same linear order $\geq$. In this setting, consider pairwise majority voting, which confronts every pair of alternatives x and y against each other, and determines that alternative x is (weakly) socially preferred to y if the number of agents who strictly prefer x to y is larger or equal to the number of agents that strictly prefer y to x. Formally, for any pair of alternatives $\{x, y\} \subset X$, we say that $x\widehat{F}\left(\succsim^1, \succsim^2, ..., \succsim^N\right) y$, which denotes "$x$ is weakly socially preferred to y", if

$$\underbrace{\#\left\{i \in N : x \succ^i y\right\}}_{\text{votes for } x} \geq \underbrace{\#\left\{i \in N : y \succ^i x\right\}}_{\text{votes for } y}$$

that is, if the number of votes for alternative x is weakly larger than those to alternative y.

We next show that single-peaked preferences avoids the presence of Condorcet cycles in social preferences. Before doing that, we define what we mean by a "median voter" which we use in our subsequent discussion and proof.

Median voter *Definition***.** Individual $m \in N$ is a median voter if

$$\#\{i \in N : x_i \geq x_m\} \geq \frac{N}{2} \text{ and } \#\{i \in N : x_m \geq x_i\} \geq \frac{N}{2}$$

That is, at least half of the population has ideal points weakly above that of individual m, and at least half of the population has ideal points weakly below that of m. A natural conclusion of this definition is that, if there are no ties in peaks (i.e., individuals with the same ideal points) and if the number of individuals is odd, then there are exactly $\frac{N-1}{2}$ individuals with ideal points strictly smaller than x_m and, similarly, $\frac{N-1}{2}$ individuals with ideal points strictly larger than x_m; ultimately implying that the median voter is unique.[13] We are now ready to provie the existence of a Condorcet winner in this setting.

Result. When preferences are single-peaked, the social preference arising from applying pair-wise majority voting has at least one alternative that cannot be defeated by any other alternative, i.e, a Condorcet winner exists.

Proof. Consider a society with N individuals, each of them exhibiting single-peaked preferences. Then, the ideal point of the median agent, x_m, cannot be defeated by majority voting by any other alternative y, i.e., x_m is a Condorcet winner. To prove this point, take any alternative $y \in X$ and suppose that the ideal point of the median agent, x_m, satisfies $x_m > y$ (the argument is analogous if we assume that $y > x_m$). For x_m to be a Condorcet winner, we then need to show that alternative x_m defeats y, that is,

$$\underbrace{\#\left\{i \in N : x_m \succ^i y\right\}}_{\text{votes to } x_m} \geq \underbrace{\#\left\{i \in N : y \succ^i x_m\right\}}_{\text{votes to } y}$$

meaning that the number of individuals who strictly prefer x_m to y is larger than the number of individuals who strictly prefer y to x_m.

Consider now the number of individuals with ideal points to the right-hand side of x_m, that is, $R_m = \#\{i \in N : x_i \geq x_m\}$. Similarly, the number of individuals with ideal points to the left of x_m can be compactly represented as $L_m = \#\{i \in N : x_m \geq x_i\}$. Therefore, $R_m \geq L_m$ and, since m is the median voter, the number of individuals with ideal points to the right of x_m must be larger than $\frac{N}{2}$, entailing

$$R_m \geq \frac{N}{2} \geq L_m.$$

We can now think again about the number of votes going to x_m and y when only these two alternatives are on the ballot. On one hand, all individuals with ideal points to the right of x_m, as captured by R_m, vote for x_m since their ideal point x_i satisfies $x_i \geq x_m > y$, weakly prefering x_m

[13] As an example, consider a setting with five voters with ideal points in $x_i \in [0, 1]$, where $x_1 < x_2 < ... < x_5$. Hence, the median voter is individual 3, with ideal point x_3, leaving the ideal points of voters 1 and 2 to the left of x_3, and the ideal points of voters 4 and 5 to the right of x_3.

to y. This is illustrated in figure 12.3.[14]

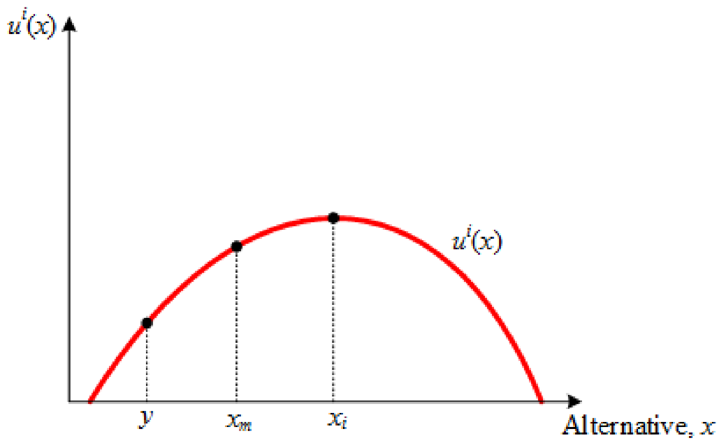

Figure 12.3. Condorcet winner with single-peaked preferences.

This means that the number of votes to x_m is weakly larger than R_m, $\#\left\{i \in N : x_m \succ^i y\right\} \geq R_m$. Inserting this result in the above inequality, we obtain

$$\#\left\{i \in N : x_m \succ^i y\right\} \geq R_m \geq \frac{N}{2} \geq L_m.$$

On the other hand, the voters to the left of x_m, as described by L_m, may prefer alternative y to x_m, but not necessarily, implying that L_m is larger than the votes going to alternative y, i.e., some voters in L_m prefer x_m to y. Therefore, $L_m \geq \#\left\{i \in N : y \succ^i x_m\right\}$, which inserted in our above inequality yields

$$\#\left\{i \in N : x_m \succ^i y\right\} \geq R_m \geq \frac{N}{2} \geq L_m \geq \#\left\{i \in N : y \succ^i x_m\right\}.$$

Focusing on the terms in the extreme left- and right-hand side of the inequality, we obtain that the number of individuals who prefer to vote for x_m is larger than those voting for y. Since alternative y is arbitrary, the median voter's ideal point, x_m, beats all other alternatives, ultimately making x_m the Condorcet winner; as required. (Q.E.D.)

As a consequence, imposing the assumption of single-peaked preferences guarantees the existence of a Condorcet winner. This is a positive result, as it helps us avoid the cyclicalities described in the Condorcet paradox. In other words, the order in which pairs of alternatives are confronted in pairwise majority voting does not affect the final outcome of the election. However, the presence of single-peaked preferences does not necessarily guarantee transitivity; as the next example illustrates.

[14] Intuitively, for every individual $i \in S$, alternative x_m is closer to i's ideal point, x_i, than alternative y is, since $x_i \geq x_m$ and $x_m > y$.

Example 12.10. Intransitive social preferences. Consider a set of three alternatives $X = \{x, y, z\}$ and $N = 4$ individuals, with the following preference profiles

$$\begin{aligned} x &\succ^1 y \succ^1 z \quad \text{for individual 1,} \\ z &\succ^2 y \succ^2 x \quad \text{for individual 2,} \\ x &\succ^3 z \succ^3 y \quad \text{for individual 3, and} \\ y &\succ^4 x \succ^4 z \quad \text{for individual 4} \end{aligned}$$

We thus have that, when we run a pairwise majority voting between alternatives x and y, we obtain

$$\#\left\{i \in N : x \succ^i y\right\} = \#\left\{i \in N : y \succ^i x\right\} = 2$$

that is, the number of individuals preferring x over y (voters 1 and 3) coincides with the number preferring y over x (voters 2 and 4). Similarly, if we confront alternatives y and z in a pairwise majority voting, we find that

$$\#\left\{i \in N : z \succ^i y\right\} = \#\left\{i \in N : z \succ^i y\right\} = 2$$

since individuals 2 and 3 vote for alternative z, while individuals 1 and 4 vote for y. Therefore, we can conclude that alternative x is socially indifferent to y and, similarly, y is socially indifferent to z. More compactly, $z\widehat{F}\left(\succsim^1, \succsim^2, \succsim^3, \succsim^4\right) y$ and $y\widehat{F}\left(\succsim^1, \succsim^2, \succsim^3, \succsim^4\right) x$. For transitivity to hold, we would need $z\widehat{F}\left(\succsim^1, \succsim^2, \succsim^3, \succsim^4\right) x$ to be satisfied. However, this result does not hold. Indeed, when alternatives z and x are presented to voters, the number of individuals preferring x to z (voters 1, 2 and 4) is larger than those preferring z to x (voter 2), that is,

$$\#\left\{i \in N : x \succ^i z\right\} = 3 \quad \text{and} \quad \#\left\{i \in N : z \succ^i x\right\} = 1$$

thus implying $x\widehat{F}\left(\succsim^1, \succsim^2, \succsim^3, \succsim^4\right) z$, which violates transitivity in the swf.

However, a Condorcet winner exists. To show that, let us run a pairwise majority voting between all pairs of alternatives, in order to test if one alternative beats all others. First, in a pairwise majority voting between x and y, there is a tie since, as described above, two individuals vote for alternative x (voters 1 and 3) and the same number of individuals vote for y (voters 2 and 4). Second, in a pairwise majority voting between y and z, there is a tie since, as discussed above, two individuals vote for alternative z (voters 2 and 3) and the same number of individuals vote for y (voters 1 and 4). Finally, in a pairwise majority voting between z and x, alternative z wins as it receives votes from voters 1, 2 and 4. Hence, alternative z becomes the Condorcet winner.[15] □

Guaranteeing transitivity in the swf. In order to guarantee that the social preference

[15] Note that a Condorcet winner, such as alternative z in this example, allows for z to either defeat all other alternatives, or produce a tie when z is confronted to some (but not all) alternatives. In other words, we cannot find another alternative that defeats z in a pairwise majority voting, thus declaring z as the Condorcet winner.

emerging from the swf is transitive, we need to impose single-peaked preferences and two additional conditions: (1) The preference relation of every individual i must be strict (that is, we no longer allow individuals to be indifferent between some alternatives); and (2) the number of individuals N is odd. While the previous example, in which the social preference was intransitive, satisfied condition (1) since individual preferences were strict, it did not satisfy condition (2), as we considered four voters. We next show that these two requirements help us obtain a transitive swf.

Result. *Consider an odd number of individuals N, each of them with strict single-peaked preferences relative to the linear order $\geq$. The social preference must be transitive.*

Proof. Consider a set of alternatives $X = \{x, y, z\}$, where

$$x\widehat{F}\left(\succsim^1, \succsim^2, ..., \succsim^N\right) y \quad \text{and} \quad y\widehat{F}\left(\succsim^1, \succsim^2, ..., \succsim^N\right) z.$$

That is, alternative x defeats y, and y defeats x. Since individual preferences are strict and N is odd, there must be one alternative in X that is not defeated by any other alternative in X. However, such alternative can be neither y (since y is defeated by x) nor z (which is defeated by y). Hence, such alternative has to be x. We can, thus, conclude that $x\widehat{F}\left(\succsim^1, \succsim^2, ..., \succsim^N\right) z$; as required to prove transitivity. (Q.E.D.)

In summary, imposing the assumptions of strict, single-peaked preferences, with an odd number of individuals, helped us guarantee not only acyclic preferences (thus producing a Condorcet winner), but also a transitive social preference relation. While this result is positive, we must recognize that our above discussion only considered that the set of alternatives X was unidimensional, i.e., $X \subset \mathbb{R}$, a segment in the real line. For instance, the alternative being considered for a vote is the percentage of government budget that a political candidate plans to spend in education. In many settings, however, candidates are evaluated on several dimensions, such as their plans for military spending, their experience, and even their looks, thus making their policy proposals multidimensional. A natural question is, then, whether we can still find a Condorcet winner in the social preference when we consider individual preferences that rank policy alternatives according to two or more dimensions. Bad news: a Condorcet winner may not exist in this setting; see Caplin and Nalebuff (1988) for a detailed analysis.[16]

3.2 Second reaction - Social welfare function

While the first reaction of the literature to Arrow's impossibility theorem restricted the set of individual preferences being considered, the second reaction allows for the intensity of individual preferences to enter into social preferences. That is, rather than seeking an ordinal measure of social preferences, we obtain a cardinal measure. In particular, this approach uses a social welfare function

$$W\left(u^1(\cdot), u^2(\cdot), ..., u^N(\cdot)\right)$$

[16] Caplin, A. and B. Nalebuff (1988) "On 64% majority rule" Econometrica, 56, pp. 787-814.

with its arguments being the utility levels of all individuals. We next describe some well-known social welfare functions, such as the utilitarian and the Rawlsian. Afterwards, we present some properties and discuss which social welfare functions satisfy each property. Formal proofs of each result are left as short end-of-chapter exercises.

3.3 Common social welfare functions

Rawlsian social welfare function. This function considers that the welfare that society obtains from an alternative x coincides with the utility level of the worst-off individual, that is,

$$W(x) = \min \left\{ u^1(x), ..., u^N(x) \right\}.$$

Utilitarian social welfare function. This function assigns an equal weight to the utility level of each individual, and it is probably the most commonly used social welfare function in economics.

$$W(x) = u^1(x) + u^2(x) + ... + u^N(x) = \sum_{i=1}^{N} u^i(x)$$

Hence, in a society with two individuals, $W = u^1(x) + u^2(x)$ which, solving for u^2, yields

$$u^2(x) = W - u^1(x),$$

thus being represented by a straight line with slope -1 in the $(u_1, u_2)-$quadrant; which can be interpreted as the "iso-welfare curve."

Example 12.11. Generalized utilitarian swf. We can also expand our previous results to the "generalized utilitarian" function of the form

$$W(x) = \sum_{i=1}^{I} \alpha^i u^i(x)$$

where $\alpha^i > 0$ represents the weight society assigns to individual i. For the case of two individuals, the generalized utilitarian swf becomes $W = \alpha^1 u^1 + \alpha^2 u^2$ which, solving for u^2, yields a social indifference curve of

$$u^2 = \frac{W}{\alpha^2} - \frac{\alpha^1}{\alpha^2} u^1,$$

thus being still a straight negatively sloped line, but its slope is now $-\frac{\alpha^1}{\alpha^2}$. Figure 12.4 depicts three social indifference curves, depending on the value of the $\frac{\alpha^1}{\alpha^2}$ ratio. In order to interpret this ratio, consider a society seeking to increase individual 1's utility in one more unit while still maintaining the welfare level unaffected (graphically represented by a rightward movement along the same social indifference curve). When society assigns a larger weight to the utility of individual 1 than 2, $\alpha^1 > \alpha^2$, the ratio becomes larger than 1 in absolute value, i.e., $-\frac{\alpha^1}{\alpha^2} > -1$, implying that the amount of u^2 that society is willing to give up (in order to increase u^1 by one unit) is relatively large. The opposite argument applies when weights satisfy $\alpha^1 < \alpha^2$, as the ratio now satisfies $-\frac{\alpha^1}{\alpha^2} < -1$, entailing that society is willing to give up a small utility from individual 2, u^2, to increase the utility of individual 1, u^1, by one unit. Finally, note that the utilitarian swf can be understood as a special case of the generalized utilitarian when weights coincide, i.e., $\alpha^1 = \alpha^2$. □

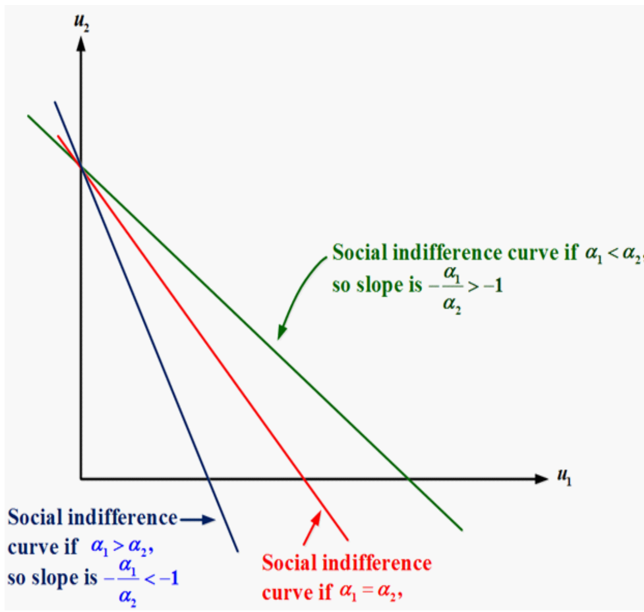

Figure 12.4. Social indifference curves for a generalized utilitarian swf.

CES social welfare function. As a summary, we can encompass all previous functional forms of social welfare functions into the following, which exhibits a familiar CES form,

$$W(x) = \sum_{i=1}^{N} \left[\left(u^i(x) \right)^{\rho} \right]^{\frac{1}{\rho}}$$

where $0 \neq \rho < 1$. Hence, the constant elasticity of social substitution between the utility of any two individuals, σ, can be expressed as $\sigma = \frac{1}{1-\rho}$. This swf satisfies a common assumption in consumer theory: strong separability. Formally, the MRS_{u^i,u^j} only depends on the utility of individuals i and j, u^i and u^j, but does not depend on the utility from any other individual $k \neq i, j$. Indeed, the MRS_{u^i,u^j} of this CES swf is

$$MRS_{u^i,u^j} = -\left(\frac{u^i}{u^j} \right)^{\rho - 1}$$

which is independent on u^k. For illustration purposes, figure 12.5 depicts three social indifference curves of a CES swf: (1) $\rho \to 1$, corresponding to linear social indifference curves, i.e., utilitarian swf; (2) $-\infty < \rho < 1$, curvy social indifference curves, resembling Cobb-Douglas indifference curves in consumer theory; and (3) $\rho \to -\infty$, corresponding to right-angled social indifference curves, i.e., Rawlsian swf.

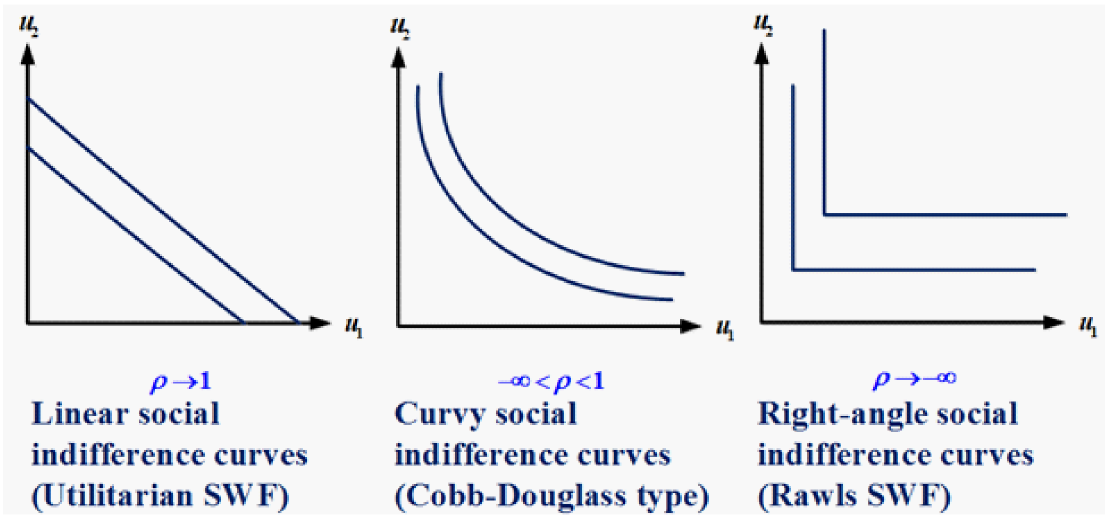

Figure 12.5. CES swf (three cases).

3.4 Social welfare functions - Properties

Since individual preferences can be represented with different utility functions, the following properties consider transformations on individual utility functions, and their effects on the social welfare function that aggregates these individual utility levels.

In particular, we are interested in guaranteeing that the social ranking of two alternatives x and y, e.g., $W\left(u^1(x), u^2(x), ..., u^I(x)\right) \geq W\left(u^1(y), u^2(y), ..., u^I(y)\right)$, is unaffected if we transform individual utility functions. Otherwise, it would be troublesome if, after applying a monotonic transformation on utility functions (which still represent the same preference relations), the social ranking between alternatives x and y changes. The next subsections investigate conditions on the monotonic transformations on utility functions that guarantee that the social ranking of alternatives is unaffected.

3.4.1 Utility-level invariance

First, consider a setting in which $u^1(x) > u^1(y)$ for individual 1, and $u^2(x) < u^2(y)$ for individual 2. In addition, assume that $u^1(y) > u^2(x)$, i.e., individual 1 is better off at his least-preferred alternative than individual 2 is. Then,

$$u^1(x) > u^1(y) > u^2(x)$$

where $u^2(y)$ must be larger than $u^2(x)$, but could rank above/below $u^1(x)$ or $u^1(y)$. Figure 12.6 provides an example of this utility ranking. First, note that individual 1 obtains a higher utility from alternative x than from y (see, respectively, points B and A in u^1); while individual 2 enjoys a higher utility from alternative y than from x (as illustrated, respectively, in points C and D in the figure). Second, point A is higher than C, thus implying that $u^1(y) > u^2(x)$, as required.

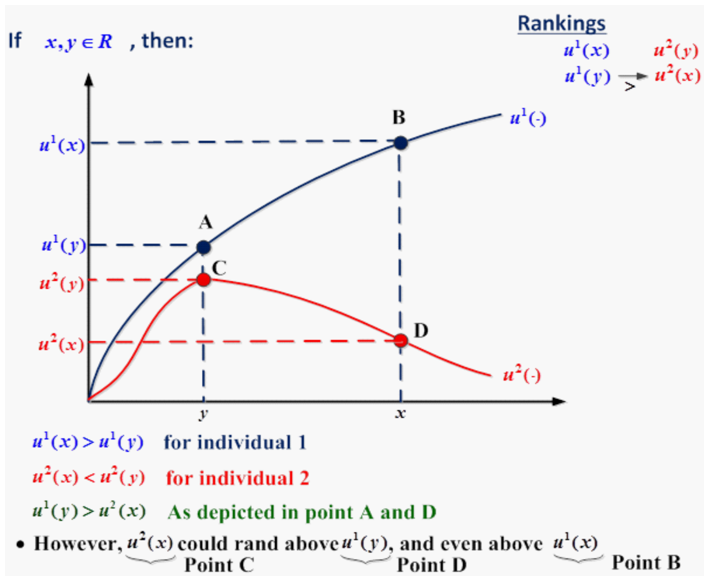

Figure 12.6. Utility-level invariance (motivation).

Assume that, in this context, society deems alternative y as socially preferred to x.[17] Now, consider strictly increasing transformations $\psi^1(\cdot)$ and $\psi^2(\cdot)$ producing the same individual ranking

$$\begin{aligned} v^1(x) &\equiv \psi^1(u^1(x)) > \psi^1(u^1(y)) \equiv v^1(y), \text{ and} \\ v^2(x) &\equiv \psi^2(u^2(x)) > \psi^2(u^2(y)) \equiv v^2(y) \end{aligned}$$

but altering the ranking across individuals, i.e., we started with $u^1(y) > u^2(x)$ but after applying these increasing transformation we obtain $v^1(y) < v^2(x)$. Hence, society would identify alternative x as socially preferred to y. However, this new social ranking is troublesome: we have not changed the individual ranking over alternatives, yet the social ranking changed! Figure 12.7 superimposes functions $v^1(\cdot)$ and $v^2(\cdot)$ on top of utility functions $u^1(\cdot)$ and $u^2(\cdot)$, showing that the individual ranking of alternatives did not change but the social ranking did.

[17] This social ranking of alternatives could be explained because society seeks to make its least well off individual as well off as possible, thus using the Rawlsian (or maxmin) criterion that we describe below. The assumption that alternative y is socially preferred to x is, however, without loss of generality; so a similar argument applies if x is socially preferred to y.

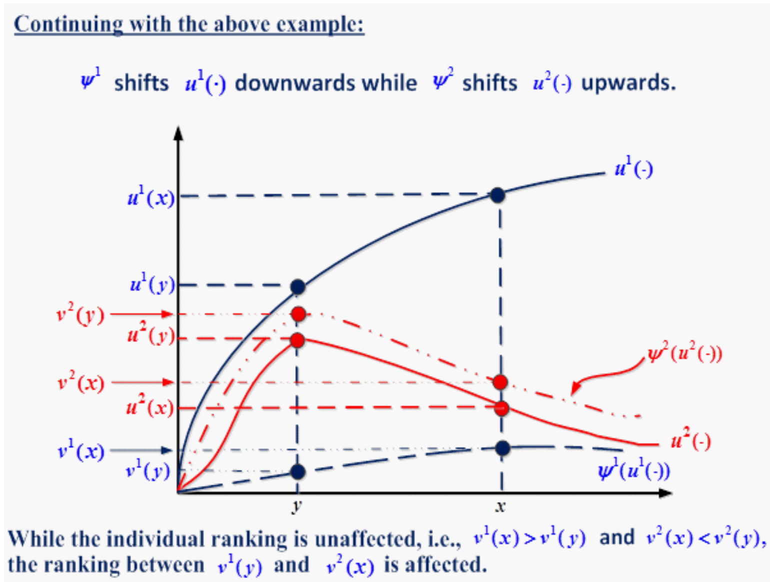

Figure 12.7. Utility-level invariance (transformations).

In order to avoid this possibility, we need to use the same monotonic transformations on both individuals' utility function, i.e., $\psi^1 = \psi^2$; as we next define.

Utility-level invariance, ULI. *Definition***.** A social welfare function $W(\cdot)$ is *utility-level invariant* if it is invariant to arbitrary, but common, strictly increasing transformations ψ applied to every individual's utility function.

This definition can alternatively be understood as follows. Consider a profile of individual preferences $\mathbf{u} \equiv \left(u^1(\cdot), u^2(\cdot), ..., u^N(\cdot)\right)$, where $\mathbf{u}(x) \equiv \left(u^1(x), u^2(x), ..., u^N(x)\right)$ and $\mathbf{u}(y) \equiv \left(u^1(y), u^2(y), ..., u^N(y)\right)$ denote the profile of individual utility levels from any two alternatives $x \neq y$. Therefore, ULI implies that

$$\text{if } W(\mathbf{u}(x)) > W(\mathbf{u}(y)) \text{ then } W(\psi(\mathbf{u}(x))) > W(\psi(\mathbf{u}(y)))$$

under a *common* strictly increasing transformation $\psi(\cdot)$. In summary, if we apply the same monotonic transformation $\psi(\cdot)$ to all individuals' utility function, the social ranking over alternatives remains unaffected.

The Rawlsian social welfare function satisifes this property holds since, after applying a common, strictly increasing, transformation ψ, on every individual's utility function, $W(x)$ varies by exactly ψ. Therefore, the ranking between any two alternatives x and y must remain unaffected, i.e., if $W(x) \geq W(y)$ then $W(\psi(x)) \geq W(\psi(y))$. We ask you to provide a more formal proof of these results in Exercise 15 at the end of the chapter.

3.4.2 Utility-difference invariance

While the utility level that each individual obtains is important in making social choices, and thus the relevance of our above discussion on ULI, another type of information often used in making social choices is related with the utility gain/loss that every individual experiences when he moves from an alternative y to another alternative x. Let us now analyze this utility difference. In particular, consider that individual 1 enjoys a utility gain $u^1(x) - u^1(y)$ when moving from y to x, while individual 2 suffers a utility loss of $u^2(x) - u^2(y)$ from the same change in alternatives, i.e., $u^1(x) - u^1(y) > 0 > u^2(x) - u^2(y)$. A common comparison is, then, whether individual 1's gain, $u^1(x) - u^1(y)$, is larger (in absolute value) than individual 2's loss, $u^2(y) - u^2(x)$, that is,

$$u^1(x) - u^1(y) > u^2(y) - u^2(x).$$

Figure 12.8a depicts a setting where individuals' utility functions are linear and this ranking holds.[18] For the swf to preserve this information, we need that monotonic transformations on $u^i(x)$ to be linear; as we next define.

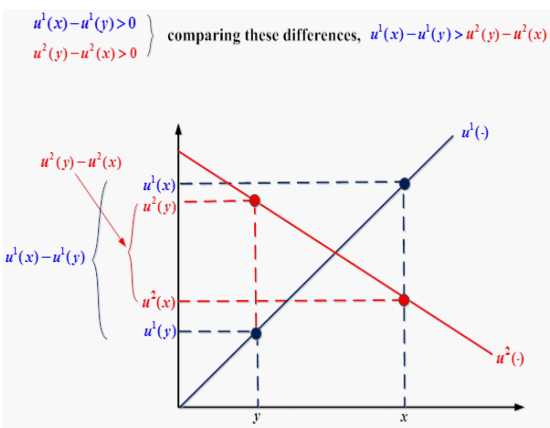

Figure 12.8a. UDI-motivation.

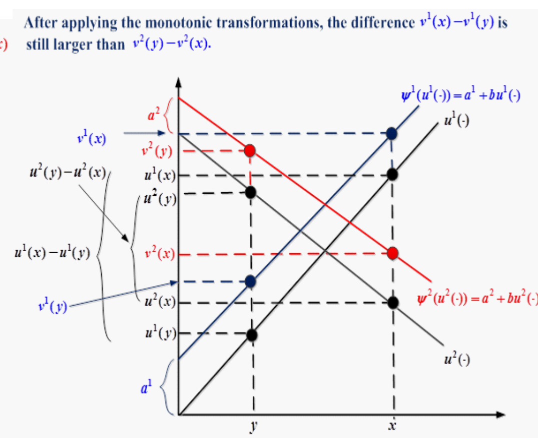

Figure 12.8b. UDI-transformations.

Utility-difference invariance, UDI. *Definition.* A social welfare function $W(\cdot)$ is *utility-difference invariant* if it is invariant to strictly increasing transformations of the following linear form

$$\psi^i\left(u^i(x)\right) = a^i + bu^i(x)$$

where $b > 0$ is common to all individuals.

While the slope of the linear transformation b coincides across all individuals, parameter a^i is

[18] Indeed, individual 1's gain of moving from alternative y to x (see arrow measuring gain $u^1(x) - u^1(y)$ in the vertical axis of figure 12.8a) offsets individual 2's loss (depicted by $u^2(y) - u^2(x)$ in the vertical axis as well).

allowed to be type-dependent. Graphically, this means that the upward or downward shift on the vertical intercept of $u^i(x)$ can vary across individuals, thus implying that the linear transformation can differ for each individual. Figure 12.7b depicts the initial utility function $u^i(x)$ and the monotonic transformation $v^i(x)$ for both individuals. The figure illustrates that, after applying a linear transformation (but not necessarily common) on both individuals' utility function, the initial ranking still holds, i.e., if $u^1(x) - u^1(y) > u^2(y) - u^2(x)$ then $v^1(x) - v^1(y) > v^2(y) - v^2(x)$ is still satisfied.[19] (The utilitarian social welfare function satisfies UDI, a proof we leave for the reader for practice in Exercise 18 at the end of the chapter.)

3.4.3 Anonymity

Anonymity. *Definition.* Let $\mathbf{u}(x)$ and $\widetilde{\mathbf{u}}(x)$ be two utility vectors of alternative x, where $\widetilde{\mathbf{u}}(x)$ has been obtained from $\mathbf{u}(x)$ after a permutation of its elements. Then,

$$W(\mathbf{u}(x)) = W(\widetilde{\mathbf{u}}(x))$$

In words, the social ranking between alternatives x and y does not depend on the identity of the individuals involved, but only on the levels of utility that each alternative produces. In a two-individual society, anonymity request that

$$W(u^1(x), u^2(x)) = W(u^2(x), u^1(x))$$

for every alternatively x, where we only permuted the identify of individuals 1 and 2.

The Rawlsian social welfare function satisfies anonymity, since the utility of the worst-off individual k, $u^k(x) = \min\left\{u^1(x), ..., u^N(x)\right\}$, regardless of how we permute the list of individuals (their identities). Similarly, the utilitarian social welfare function $W(\mathbf{u}(x)) = \sum_i u^i(x)$ satisfies this property since the result of the sum of utility levels is unaffected if we "shuffle" individuals' identities. However, the generalized utilitarian social welfare function, $W(\mathbf{u}(x)) = \sum_i \beta^i u^i(x)$, where $\beta^i \in [0, 1]$ represents the weight that society assigns to the utility of individual i, violates anonymity since total social welfare may increase or decrease after we shuffle individuals' identities.

3.4.4 Hammond Equity

Hammond Equity, HE. *Definition.* Let $\mathbf{u}(x)$ and $\mathbf{u}(y)$ be the utility vectors of two distinct alternatives x and y, where $u^k(x) = u^k(y)$ for every individual k except for two individuals, i and j. If

$$u^i(x) < u^i(y) < u^j(y) < u^j(x)$$

then $W(\mathbf{u}(y)) \geq W(\mathbf{u}(x))$.

[19] As a remark, note that parameter a^i is allowed to be positive, as in figure 12.8b whereby u^i experiences an upward shift, or negative, where u^i would shift downwards.

Intuitively, HE says that society has a preference towards the alternative that produces the smallest variance in utilities across individuals (which corresponds to alternative y in the above definition). Figure 12.9 depicts, for a society with two individuals, the inequality which constitutes the premise for HE. Since this inequality implies that alternative y produces a utility pair that lies closer to the 45-degree line, alternative y is then associated to more equality than x. According to HE, the more equal alternative y generates a larger social welfare than x.

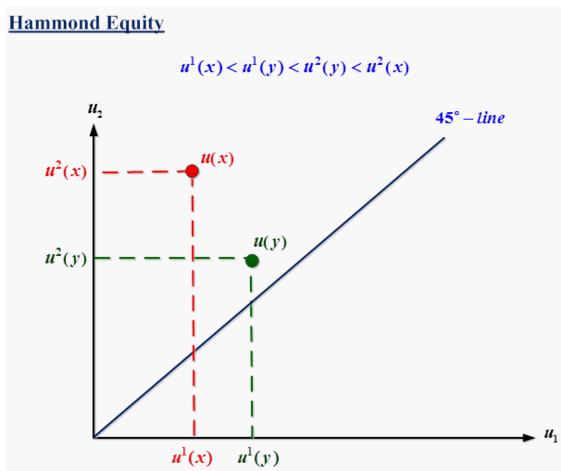

Figure 12.9. Hammond Equity.

While HE seems a reasonable property, it is often criticized because it focuses on equity, but potentially ignores the sum of utility levels that individuals obtain (i.e., the size of the pie). Consider, for instance, the following utility levels, which satisfy the premise in HE

$$u^i(x) = 1 < u^i(y) = 1.1 < u^j(y) = 1.2 < u^j(x) = 100.$$

A social welfare function satisfying HE prefers alternative y to x, as it is more equal. However, the sum of utility levels under alternative y is only $1.1 + 1.2 = 2.3$, being much smaller than the sum of utilities with alternative x, $1 + 100 = 101$.

It is straightforward to show that the Rawlsian social welfare function satisfies HE. First, let $u^k(x) = u^k(y) = \overline{u}$, which can lie in any of the regions $A - E$ in figure 12.10. We next demonstrate that the Rawlsian social welfare function produces $W(x) \leq W(y)$ regardless of the specific region where $\overline{u}$ lies.

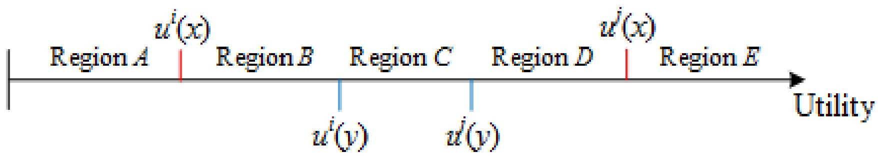

Figure 12.10. $u^k(x) = u^k(y)$ can lie in regions $A - E$.

In Region A, we obtain $W(x) = u^k(x)$ since $u^k(x)$ is lower than $u^i(x)$ and $u^j(x)$, which coincides with $W(y) = u^k(y)$ since $u^k(x) = u^k(y)$ by assumption. In Region B, we find that $W(x) = u^i(x)$ since $u^k(x)$ is lower than $u^j(x)$ and $u^k(x)$ for all regions $B - E$, and $W(y) = u^k(y)$, entailing $W(x) < W(y)$ in this case. In regions $C - E$, we have $W(x) = u^i(x)$ and $W(y) = u^i(y)$, which implies that $W(x) < W(y)$ since $u^i(x) < u^i(y)$ by assumption. Therefore, in all regions, society ranks alternative y as weakly preferred to x.

The utilitarian social welfare function, however, violates HE; as illustrated in the above numerical example where the sum of utilities was 101 with alternative x and only 2.3 with alternative y. A similar argument applies to the generalized utilitarian social welfare function.

4 Alternatives to majority voting

In previous sections, we criticized both majority voting, as it could lead to ciclycalities and agenda manipulation, and the Condorcet criterion, as it could lead to no candidate being selected as the winner. A natural question is, then, whether other voting procedures, especially those commonly observed in real life elections, produce Condorcet winners. In this section, we finish this chapter taking a relatively applied approach, first describing voting procedures and then comparing them in terms of their properties.

4.1 A list of voting procedures

We first describe two familiar voting procedures: majority rule, and the Condorcet winner.

Majority rule: Society chooses the candidate who is ranked first by more than half of the voters.

Condorcet criterion: Society chooses the candidate who defeats all others in pairwise elections using majority rule.

Example 12.11. Applying majority rule and the Condorcet criterion. Consider three candidates, A, B, and C; and three voters who rank these candidates as follows:

Voter 1	Voter 2	Voter 3
A	A	C
B	C	B
C	B	A

Table 12.9. Applying majority rule and the Condorcet criterion.

For instance, voter 2 ranks candidate A at the top of his list, candidate C next, and candidate B last. If majority voting is used, and each voter cast his ballot for his most preferred candidate, candidate A would receive two votes (from voters 1 and 2), candidate B would receive no votes, and candidate C would receive one vote (from voter 3). Hence, candidate A would be the winner according to majority rule.

If, instead, the Condorcet criterion was used, a pairwise vote between candidates A and B would yield A as the winner (since A is preferred to B by voters 1 and 2, while B is preferred to A by only voter 3). The winner of this pairwise confrontation, candidate A, would then be paired against the remaining candidate, candidate C, still yielding candidate A as the winner (in this case, A is preferred to C by voters 1 and 2, while only voter 3 prefers C to A). As a consequence, A would be the candidate winning the election according to pairwise majority voting, thus becoming the Condorcet winner. Therefore, the winner according to majority rule and the Condorcet criterion coincide. This coincidence in outcomes occurs regardless of the number of candidates, the number of voters, and their preferences. Other voting methods, however, do not necessarily produce a Condorcet winner as we show below. □

Majority rule with runoff election: If one of the m candidates receives more than half of the votes, then he/she is the winner. Otherwise, a second (runoff) election is held between the two candidates receiving the most votes on the first ballot. The candidate receiving the most votes on the second election is declared the winner. Runoff elections are common in several countries, such as France, Russia, Chile, Argentina, and Brazil.

Example 12.12. Applying majority rule with runoff election. Consider four candidates running for office $\{X, Y, Z, W\}$, and five voters, 1-5, with the following preferences

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
X	X	Y	W	W
Y	Y	Z	Y	Y
Z	Z	W	Z	Z
W	W	X	X	X

Table 12.10. Applying majority rule with runoff election.

If society uses majority rule with runoff election, candidate X receives two votes (from voters 1 and 2), candidate Y receives only one vote from voter 3, candidate Z receives no votes, and candidate W receives two votes from voters 4 and 5. Hence, no candidate receives more than 50% of the votes (since four candidates run for office in this example, we need at least three votes going to the same candidate for her to win). A runoff election is then held between the two candidates receiving the most votes on the first ballot: candidates X and W. In this runoff election, voters 1 and 2 cast their ballot for candidate X (as they prefer X to W), while voters 3-5 vote for candidate W, thus making W the winner. As a practice, note that, if this society uses majority rule (without runoff election) candidates X and W would receive two votes each, thus producing a tie. □

Plurality rule: Voters choose one candidate in the ballot, and the candidate with most votes wins.

This voting method is often known as "first past the post" since the candidate accumulating the most votes wins, even if he/she receives less than half of total votes.

Example 12.13. Applying plurality rule. Consider the preference profile in Example 12.11 again. In such a setting, voters 1 and 2 rank candidate A at the top of their list, while voter 3 ranks candidate C. (Candidate B is not ranked first by any voter.) Hence, candidate A is voted by most voters, and is declared the winner under plurality rule. □

Instant runoff: Every voter ranks candidates in order of preference. If a candidate is ranked the highest by more than 50% of voters, she is declared the winner. Otherwise, the candidate ranked the highest by the fewest voters is eliminated. The top remaining choices in every ballot are then counted again, eliminating the candidate ranked as the top choice by the fewest voters. The process is repeated until only one candidate remains as the top choice for a majority of voters (more than 50%). When the number of candidates reduces to two, the above voting system becomes an instant runoff since it allows for a comparison of the two top candidates head-to-head.

Intuitively, instant runoff uses voters' ranked choice ballots to simulate a traditional runoff in a single round of voting, rather than asking voters to cast their ballots several times. Since voters are asked to provide their ranking of candidates, instant runoff is often known as "ranked-choice voting," or "preferential voting." Instant-runoff is used to elect members of the Australian House of Representatives, the President of Ireland, members of Congress in Maine (U.S.), in local elections in several countries[20] and, as a curiosity, to select the Academy Award for Best Picture.

Example 12.14. Instant runoff and the "spoiler" effect. Consider the preference profile in Table 12.11, where 35% of voters prefer candidate A to B and B to C, 25% prefer candidate B to A and A to C, and the remaining 40% of the voters prefer candidate C to B and B to A.

[20] Examples include London (United Kingdom), Wellington (New Zealand), Minneapolis and St. Paul in Minnesota, San Francisco and Oakland in California, and Portland in Maine.

Intuitively, the first two groups of voters (who account together for 60% of votes) regard candidates A and B are relatively similar, while C is perceived as an extremist. If the election used plurality voting, candidate A would receive 35% of the votes, candidate B would receive 25%, and C would achieve 40% of the votes, thus being declared the winner. This can be regarded as a surprising outcome of the election, since 60% of electors prefer both candidates A and B over C.[21]

35%	25%	40%
A	B	C
B	A	B
C	C	A

Table 12.11. Applying instant runoff-I.

In contrast, with instant runoff, we first eliminate candidate B since it was ranked as the top candidate by the fewest voters (25%), which yields Table 12.12. We now count how many voters ranked candidate A as the top choice ($35 + 25 = 60\%$) against those ranking candidate C as the top choice (40%), which implies that candidate A wins the election.

35%	25%	40%
A	A	C
C	C	A

Table 12.12. Applying instant runoff-II.

In this example, candidate B is often regarded as the election "spoiler" when plurality rule is used, since his presence changes the outcome of the election. (To see this, note that if candidate B did not run for office, the relevant table of voter preferences should be Table 12.12, where candidate A wins according to both plurality rule and instant runoff.) Generally, the spoiler effect occurs when the presence of a candidate that voters regard as similar to other candidate (the spoiler) splits the vote, allowing a third candidate to win. Examples include the 2000 U.S. presidential election between Bush, Gore, and Nader, or the 2002 French presidential election between Chirac, Jospin, and LePen. □

While instant runoff reduces the chances of spoiler effects, this can still exist. Other voting methods we discuss below, such as approval voting, completely eliminate the risk of an spoiler. Mathematically, the spoiler effect indicates that the swf aggregating individual preferences violates the IIA property, since the ranking between candidates A and B does not change when we move from Table 12.12 to 12.11, where the only thing that changed was the addition of candidate B. However, under plurality voting, the outcome of the election (i.e., the social ranking of candidates)

[21] This outcome may also give rise to "strategic voting," since voters prefering A and B to C in the first two columns of Table 12.11 may have incentives to choose a candidate different from their top choice to reduce the chances of C winning. For instance, the voters with preferences $B \succ A \succ C$ in the second column (which only account for 25% of the votes) may vote for their second-best candidate A since that could secure A wins the election to C.

changes from A being the winner in Table 12.12 to C being the winner in Table 12.11. Hence, the presence of a new alternative (the spoiler candidate B) alters the social ranking, thus violating IIA. In our example, however, instant runoff produces the same winning candidate, thus not violating IIA.

The Hare system: First, each voter indicates the candidate he ranks highest of the m candidates. Second, remove from the list the candidate ranked the highest by the fewest number of voters. Third, repeat the procedure for the remaining $m-1$ candidates. Continue until only one candidate remains in the list, who is declared the winner.

Example 12.15. Applying the Hare system. Consider again the preference profile in Example 12.11. As we discussed in Example 12.13, candidate A is ranked highest by two voters, C is ranked highest by one voter, while B is not ranked highest by any voter. Hence, candidate B is removed from the list. Once candidate B is removed from the list, every voter is asked to rank the remaining candidates, A and C, which yields the preference profile in Table 12.13.

Voter 1	Voter 2	Voter 3
A	A	C
C	C	A

Table 12.13. Applying the Hare system.

Candidate A is now ranked highest by two voters (voter 1 and 2), C is ranked highest by one voter (voter 3), which implies that candidate C is removed from the list. Therefore, A is the only candidate remaining, and he is declared the winner. As a practice, note that the winner according to the Hare system coincides with that under plurality voting identified in Example 12.13. □

Variations of the Hare system are used in elections in Australia and Ireland. While the Hare system is often proposed as an alternative to the plurality voting system common in most developed countries, it still suffers from two problems: (1) it can fail to select the Condorcet winner (even if it exists); and (2) it violates monotonicity. (We ask you to show these to points in Exercise 6 at the end of the chapter.)

The Coombs system: This voting procedure can be understood as the opposite of the Hare system. (To emphasize the differences, the next description italicizes the words that changed relative to the Hare system.) First, each voter indicates the candidate he ranks *lowest* of the m candidates. Second, remove from the list the candidate ranked the *lowest* by *most* voters. Third, repeat the procedure for the remaining $m-1$ candidates. Continue until only one candidate remains in the list, who is declared the winner.

Example 12.16. Applying the Coombs system. Consider three candidates, A, B, and C;

and three voters who rank these candidates as follows:

Voter 1	Voter 2	Voter 3
A	A	C
B	C	A
C	B	B

Table 12.14. Applying the Coombs system.-I

(This preference profile is similar to that in Example 12.11 but with a twist in voter 3's preferences for candidates A and B.) In this context, candidate B is ranked the lowest by two voters (voters 2 and 3), C is ranked the lowest by only voter 1, while A is not ranked the lowest by any voter. We can then proceed to remove candidate B from the list, as illustrated in the table.

Voter 1	Voter 2	Voter 3
A	A	C
C	C	A

Table 12.15. Applying the Coombs system.-II

We can now identify which candidate is ranked the lowest by most voters. In particular, candidate C is ranked the lowest by voters 1 and 2, while A is ranked the lowest only by voter 3. Hence, candidate C is removed from the list, leaving candidate A as the only surviving candidate, who is declared the winner according to the Coombs system; a result that coincides with the winner under plurality rule and the Hare system. □

The Borda count: First, each voter gives a score $s \in [1, m]$ to each of the m candidates, i.e., he gives m points to his most preferred candidate, $m - 1$ points to the second most preferred candidate, ..., and one point to his least preferred candidate. The candidate receiving the highest number of points is declared the winner.

This approach to aggregate preferences is rarely used in elections, but it is relatively common in college sports, such as ranking NCAA teams in the US (especially famous for college basketball), or identifying the most valuable player (MVP) in sport tournaments.

Example 12.17. Applying Borda count. Consider three candidates A, B and C; and three voters who are asked to score each candidate with a number 1-3. In this setting, a ballot would ask:

> "Please give a score 1-3 to each of the three candidates in the following list, writing 3 next to your most preferred candidate, 2 next to your second most preferred candidate, and 1 next to your least preferred candidate."

Here are examples of possible ballots marked by voters 1 and 2.

Voter 1		Voter 2		Total points	
A	3	A	1	A	$3+1=4$
B	2	B	3	B	$2+3=5$
C	1	C	2	C	$1+2=3$

Table 12.16. Applying the Borda count

Intuitively, candidate A is the most preferred by voter 1, followed by candidate B, and candidate C. In contrast, voter 2 prefers candidate B, followed by C, and ultimately by A. In this context, candidate A receives a total of 4 points, B receives 5 points, and C receives only 3 points, implying that B is declared the winner under Borda count. □

Approval voting: First, each voter votes for the k candidates he ranks highest of the m candidates, where k can vary from voter to voter and $k \in (1, m)$. The candidate with the most votes is declared the winner.

Approval voting is mostly used in elections for private associations, such as Mathematical Association of America and the American Mathematical Society, the selection of the Secretary-General of the United Nations, papal conclaves, and elections in 19th century England.

Example 12.18. Applying approval voting. Consider four candidates A, B, C, and D; and three voters who are asked to vote for one, two or three candidates, i.e., $k \in (1, 4)$ which implies that the number of votes, k, must be either $k = 2$ or $k = 3$. In such a setting, a ballot's instructions would read:

> "In the next list of three candidates, please mark a cross next to the candidate (or candidates) you want to vote for. You can mark a cross next to two or three candidates."

Examples of ballots marked by voters 1-4 could look like the following:

Voter 1		Voter 2		Voter 3		Voter 4	
A	X	A		A		A	X
B	X	B	X	B	X	B	
C		C	X	C	X	C	
D		D		D	X	D	X

Table 12.17. Applying approval voting.T

Table 12.17 indicates that voter 1 deems candidates A and B as acceptable, while C and D is regarded as unacceptable.[22] Approval voting would then sum the number of votes that each

[22] Voter 2 deems candidates B and C as acceptable, but A and D are unacceptable. A similar intuition applies to the votes from voters 3 and 4.

candidate receives: candidate A receives two votes (one from voter 1 and another from voter 4), B receives three votes (from voters 1-3), C receives two votes (from voters 2 and 3), and D obtains two votes (from voters 3 and 4). Since B is the candidate receiving the most votes, he is declared the winner according to approval voting. $\square$

4.2 Evaluating voting procedures

Given the different voting procedures suggested above (and other variations we could easily construct), which criteria can we use to compare them? We next briefly describe two common criteria.

Decisiveness. *Definition.* The voting procedure picks a winner.

As discussed in previous sections, when the number of candidates is only two, $m = 2$, all voting procedures are decisive. However, when $m > 2$, majority voting and the Condorcet criterion are not necessarily decisive, but all other voting procedures are decisive.

Condorcet winner. *Definition.* The winner of the voting procedure coincides with the Condorcet winner.

While most voting procedures identify a winner, i.e., they are decisive, such a winner doesn't need to coincide with the Condorcet winner; as the following example illustrates. Therefore, majority rule selects a Condorcet winner (if one exists), but all other voting procedures may select a winner that is not necessarily the Condorcet winner (even when one exists).

Example 12.19. Winner does not need to be the Condorcet winner. Consider four candidates running for office, $\{X, Y, Z, W\}$, and five voters, 1-5, with the following preference ranking.

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
X	X	Y	Z	W
Y	Y	Z	Y	Y
Z	Z	W	W	Z
W	W	X	X	X

Table 12.18. Plurality voting does not produce a Condorcet winner.

Using plurality voting, candidate X is ranked the highest by two voters (1 and 2), while candidates Y, Z, and W are ranked highest by only one voter each. Hence, candidate X is ranked highest by the largest number of voters, and becomes the winner according to plurality voting.[23]

However, X is not the Condorcet winner. Indeed, if candidates X and Y are confronted in a pairwise vote, Y wins as Y is preferred to X by three out of five voters (i.e., voters 3-5). The winner of this pairwise election, candidate Y, is then confronted against Z in a pairwise vote, where Y

[23] As a practice, find the winner when other voting procedures are used, such as the Borda count.

wins again since Y is preferred to Z by four voters (1, 2, 3 and 5). Finally, Y faces the remaining candidate W, which yields Y to be the winner again, as Y is preferred to W by four voters (1-4). In summary, candidate Y beats all other candidates in pairwise votes, and thus becomes the Condorcet winner; which does not coincide with the winner according to plurality voting. $\square$

5 Appendix 12.A - Unifying social welfare functions

In previous sections, we described the properties of different swfs, but did not discuss how societies choose one swf over another. The literature has mainly considered two approaches: Harsanyi's and Rawls'. Both approaches assume that individuals do not yet know which position they will occupy in society. That is, before being borned, every individual i cannot perfectly anticipate his utility level u^i, thus not knowing whether he will be one of the individuals with the highest or lowest utility level in society. This assumption by both approaches is commonly referred to as individuals' "veil of ignorance," after Rawls (1971). However, the two approaches differ in how every individual i assigns probabilities to each of the possible positions he could occupy, as we discuss below. While both approaches seem at first glance incompatible, we show below that they can actually be modeled as special cases of a more general ("unified") approach.

Harsanyi's approach. Harsanyi claims that individuals assign an *equal probability* to the prospect of being in any possible position in society, which is often referred to as the "principle of insufficient reason." Hence, if there are N individuals in a society, there is a probability $\frac{1}{N}$ that individual i will end up in the position of any of these N individuals, yielding a utility $u^i(x)$, thus implying that i's expected utility is

$$\sum_{i=1}^{N} \frac{1}{N} u^i(x)$$

Therefore, when society chooses between two alternatives x and y, alternative x is socially preferred if

$$\sum_{i=1}^{N} \frac{1}{N} u^i(x) > \sum_{i=1}^{N} \frac{1}{N} u^i(y) \Longleftrightarrow \sum_{i=1}^{N} u^i(x) > \sum_{i=1}^{N} u^i(x)$$

which exactly coincides with the condition provided by the utilitarian swf. This explains why Harsanyi's approach is often used to support the use of utilitarian swfs.

Rawls' approach. In contrast, Rawls claims that individuals have no empirical basis for assigning probabilities to each position, whether equal or unequal probabilities to each position.[24] Assuming people are risk averse, he argues that individuals would order alternatives according to which one provides him with the highest utility in case he ends up as society's worst-off member.

[24] That is, Rawls viewed every individual i's initial assessment in a setting of complete ignorance.

Thus, alternative x is socially preferred to y if and only if

$$\min\left\{u^1(x), ..., u^N(x)\right\} \geq \min\left\{u^1(y), ..., u^N(y)\right\}$$

which is a purely maximin criterion, i.e., society should choose the alternative that maximizes the utility of the worst-off individual.

Unification of both approaches. Let us now show that Rawls' and Harsanyi's approaches can be modeled as special cases of a more general setting. First, take a utility function $u^i(x)$ for individual i. Since the underlying preferences of this individual can also be represented by monotonic transformations of $u^i(x)$, we can apply the following concave transformation to $u^i(x)$

$$v^i(x) \equiv -u^i(x)^{-a}, \ \text{ where } a > 0$$

where $v^i(x)$ can be understood as the vNM utility function of this individual, with parameter a capturing his degree of risk aversion; as Figure 12.11 depicts.[25]

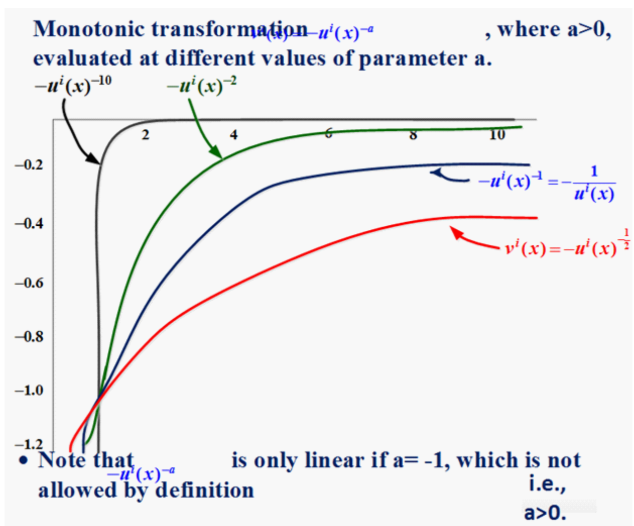

Figure 12.11. Concave transformation of $u^i(x)$.

Using the Harsanyi's approach on this monotonic transformation, yields a social welfare function

$$W = \sum_{i=1}^{N} v^i(x) \equiv -\sum_{i=1}^{N} -u^i(x)^{-a}$$

Importantly, the social ranking of alternatives provided by this swf must coincide with that of

[25] Note that function $v^i(x) \equiv -u^i(x)^{-a}$ is only linear if $a = -1$, which is not allowed since by definition parameter a satisfies $a > 0$. In addition, when a decreases, function $v^i(x)$ becomes more concave, thus reflecting a higher degree of risk aversion.

its monotonic transformation W^*, where $W^* \equiv (-W)^{-\frac{1}{a}}$, entailing

$$W^* \equiv (-W)^{-\frac{1}{a}} = \left(-\sum_{i=1}^{N} -u^i(x)^{-a}\right)^{-\frac{1}{a}} = \left(\sum_{i=1}^{N} -u^i(x)^{-a}\right)^{-\frac{1}{a}}$$

Hence, if we relabel parameter a as $-a = \rho$, we can express swf W^* as

$$W^* = \left(\sum_{i=1}^{N} -u^i(x)^{\rho}\right)^{\frac{1}{\rho}}$$

which coincides with the CES swf described in Section 12.3.3. Therefore, when $\rho \to -\infty$, the parameter of risk aversion a becomes $a \to \infty$, and the above swf approaches the maximin criterion by Rawls as a limiting case. In addition, the Rawlsian criterion becomes a special case of Harsanyi's approach when individuals become infinitely risk averse. Finally, when $-\infty < \rho < 1$, the parameter of risk aversion a satisfies $a \in [0, +\infty)$. (Note that we do not claim $a > -1$ since $a > 0$ by definition.) In that scenario, individuals are risk averse (but not infinitely), and social indifference curves are curvy.

6 End-of-Chapter Exercises

1. **Majority voting - Some properties** Consider majority voting between two alternatives x and y, so the preferences of every individual i over these two alternatives can be represented as $\alpha_i = \{1, 0, -1\}$, where $\alpha_i = 1$ indicates that individual i strictly prefers x to y; $\alpha_i = 0$ reflects that he is indifferent between alternatives x and y; and $\alpha_i = -1$ represents that he strictly prefers y to x. Let us check that, in this context, majority voting satisfies the following three properties: (a) symmetry among agents, (b) neutrality between alternatives, and (c) positive responsiveness.

2. **Three examples of social welfare functionals** In this exercise, we consider a setting with two alternatives x and y, and discuss three specific social welfare functionals $F(\alpha_1, .., \alpha_I)$ in parts (a)-(c) below. For each functional, determine whether or not it satisfies the three properties of majority voting (symmetry among agents, neutrality between alternatives, and positive responsiveness).

   (a) Let us first consider the lexicographic social welfare functional

$$F(\alpha_1, .., \alpha_N) \begin{cases} \alpha_1 \text{ if } \alpha_1 \neq 0 \\ \alpha_2 \text{ if } \alpha_1 = 0 \text{ and } \alpha_2 \neq 0 \\ \alpha_3 \text{ if } \alpha_1 = \alpha_2 = 0 \text{ and } \alpha_3 \neq 0 \\ ... \end{cases}$$

   Intuitively, society selects the alternative that individual 1 strictly prefers. However, if he is indifferent between alternatives x and y, society follows the strict preferences of individual 2 (if he has a strict preference over x or y). If both individuals 1 and 2 are indifferent between x and y, the strict preferences of individual 3 are considered, and so on.

(b) A constant social welfare functional $F(\alpha_1, ..., \alpha_N) = 1$ for all $(\alpha_1, ..., \alpha_N)$, thus representing that society chooses alternative x over y regardless of the profile of individual preferences $(\alpha_1, ..., \alpha_N)$.

(c) A constant social welfare functional $F(\alpha_1, ..., \alpha_N) = 0$ for all $(\alpha_1, ..., \alpha_N)$, thus indicating that society is indifferent between alternatives x and y regardless of the profile of individual preferences $(\alpha_1, ..., \alpha_N)$.

3. **An alternative proof of Arrow's impossibility theorem, based on Geanakoplos (2005).**[26] Genakopolos (2005) provides three proofs of Arrow's imposibility theorem. The first one, probably the most graphical, was already discussed in the chapter. In this exercise, we focus on the third proof (see pages 214-215 of the article). This proof shows that Arrow's axioms imply "strict neutrality," which we can informally interpret as that every decision must be independent of the names of the alternatives (anonymity). We use strict neutrality to show that any pivotal voter is a dictator.

   Consider the following preference relation.

$$\begin{aligned} a \succ^i b \quad &\text{for all } i \in S, \text{ but} \\ b \succ^i a \quad &\text{for all } i \in I \backslash S. \end{aligned}$$

   where group S is a subset of individuals in a finite population I. Intuitively, all individuals in group S strictly prefer alternative a to b, while all other individuals in I strictly prefer alternative b to a.

   (a) *Strict neutrality.* Show that if the social preference relation coincides with the preference relation of group S over alternatives (a, b), then the preference relation of group S over any alternatives (α, β) also determines the social preference relation over alternatives (α, β).

   (b) *A pivotal voter exists, and he is a dictator.* Show that there exists a dictator in group S. (*Hint*: Assume $b \succ^i a$ for all $i \in I$. Begin with $i = 1$, successively move a above b for every individual in S, and then find an individual n^* whose move changes the social preference relation from $b \succ^i a$ to $a \succ^i b$ for all $i \in I$.)

4. **An Alternative Proof to Arrow's Impossibility Theorem, based on Maskin and Sen (2014).**[27] In this exercise, we examine an alternative proof to Arrow's Impossibility

[26] Geanakoplos, J. (2005) "Three brief proofs of Arrow's Impossibility Theorem," Economic Theory, 26, pp. 211-215.
[27] Maskin E. and Sen A. (2014). *The Arrow Impossibility Theorem*, pp. 33-37. New York, NY: Columbia University Press.

Theorem. Following our discussion in this chapter, we consider sets of at least three alternatives (e.g., candidates running for office) and four assumptions on social welfare functions: unrestricted domain (U), independence of irrelevant alternatives (IIA), Pareto principle (P), and no dictatorship (ND). Like in the proof presented in the chapter, we seek to show that any social welfare function F which satisfies U, IIA and P, violates ND. For simplicity, we split the proof in different steps.

(a) *Spread of Decisiveness.* Show that if a subset of agents $S \subset N$ is decisive over a particular set of alternatives $\{x, y\}$, then S is globally decisive over any set of alternatives $\{w, z\} \in \mathcal{X}$.[28]

(b) *Contraction of Decisive Sets.* Show that if a set of individuals $T \subset N$ is decisive, then some partition of T is also decisive. [*Hint*: Consider two disjoint partitions of set T, T_1 and T_2, where $T_1 \cap T_2 = \varnothing$ and $T_1 \cup T_2 = T$, allowing for the preference orderings of the two groups differ.]

(c) Show that Arrow's Impossibility Theorem holds.

5. **Plurality rule and independent candidates with few supporters** Consider the 2000 US Presidential elections between George W. Bush (B), Al Gore (G), and Ralph Nader (N). Assume that voter preferences are as follows

Bush supporters, 46%	Gore supporters, 45%	Nader supporters, 9%
B	G	N
G	B	G
N	N	B

In words, the first column indicates that 46% of voters are Bush supporters, valuing him above Gore, and Gore above Nader. The second column represents that 45% of voters were Gore supporters, valuing him above Bush, and Bush above Nader. Finally, the last column illustrates that only 9% of voters were Nader supporters, valuing him above Gore, and Gore above Bush. Nader was considered the independent candidate, since it did not represent the Republican or Democratic party.

(a) Show that Gore is the Condorcet winner.

(b) Show that Bush wins the election by plurality rule.

(c) Show that if Nader did not run, Gore would have won the election by plurality rule.

(d) Repeat your analysis of parts (a)-(c) assuming now that Gore supporters (the second column) preferences change to $G \succ N \succ B$, thus moving Nader to a "second best" position rather than their last preferred candidate. How are the above results affected by this preference change?

[28] For a general proof with three alternatives, for instance, $\{x, y, z\} \in \mathcal{X}$, refer to step 1 of Proposition 21.C.1 in Mas-Colell, Whinston, and Green (1995, p.797).

6. **Strategic voting under majority rule.** Consider the following three group of voters (A, B and C) with their corresponding ranking of preferenced candidates (x, y and z).

A	B	C
0.35	0.33	0.32
x	y	z
y	z	x
z	x	y

   (a) Show that majority voting yields to a Condorcet cycle if every individual votes for his most preferred candidate.

   (b) Show that if some voters in group A vote for their second-best candidate y, rather than their first-best candidate x, they guarantee that their least prefreed candidate z does not win. This will prove that majority rule is manipulable, that is, every voter does not necessarily vote for his most preferred alternative. An example suffices.

7. **Hare system - Two problems** Consider three candidates running for office, X, Y, Z, and 17 voters with the following preference ranking.

Voters 1-7	Voters 8-12	Voters 13-17
X	Y	Z
Y	Z	Y
Z	X	X

   (a) Show that candidate X wins according to the Hare system.

   (b) Show that Y is the Condorcet winner.

8. **Majority loser criterion.** Given that all voting methods present problems under some preference orderings, an alternative could be to start ruling out candidates in an election. For instance, we could rule out the candidate ranked the lowest by more than half of the voters, which is referred as the "Majority Loser". This approach, however, can also be problematic, since we may eliminate candidates who would become the winner according to some of the voting methods described in the chapter. To illustrate this possibility, consider a setting with three candidates $A - C$, and 9 voters, with the following preferences.

4 voters	3 voters	2 voters
A	B	C
B	C	B
C	A	A

   (a) Find the winner according to plurality rule.

(b) Eliminate the majority loser.

9. **Sequential pairwise voting** Consider an alternative approach to aggregate individual preferences, known as "sequential pairwise voting." In this voting procedure, every pair of alternatives is compared by majority rule; the winner is then compared against another alternative; and so on until we exhaust all possible alternatives. Consider that four candidates run for office $\{A, B, C, D\}$. We then compare candidates A and B by majority rule, the winner is compared against candidate C, and the winner is finally compared against candidate D. In this exercise, we seek to show that sequential pairwise voting may lead to outcomes that are not Pareto optimal.

   (a) Consider that voter preferences are given by

Voter 1	Voter 2	Voter 3
A	B	A
B	D	C
C	C	B
D	A	D

   Find the candidate that wins the election by sequential pairwise voting. Is the outcome of the election Pareto optimal?

   (b) Consider now that voter preferences change to

Voter 1	Voter 2	Voter 3
A	B	C
B	D	A
C	C	B
D	A	D

   Find the candidate that wins the election by sequential pairwise voting. Show that the outcome of the election is *not* Pareto optimal?

10. **Borda count and Condorcet winner.** Consider three candidates running for office, $\{X, Y, Z\}$, and five voters, 1-5, with the following preference ranking.

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
X	X	X	Y	Y
Y	Y	Y	Z	Z
Z	Z	Z	X	X

   (a) Find the Condorcet winner.

   (b) Find the winner according to the Borda count.

11. **Hare system and Condorcet winner.** Consider four candidates running for office, $\{X, Y, Z, W\}$, and five voters, 1-5, with the following preference ranking.

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
Y	W	X	Y	W
X	Z	Z	Z	X
Z	X	W	X	Z
W	Y	Y	W	Y

(a) Find the Condorcet winner.

(b) Find the winner according to the Hare system.

12. **Coombs system and Condorcet winner.** Consider four candidates running for office, $\{A, B, C, D\}$, and 21 voters with the following preference ranking.

9 voters	6 voters	4 voters	1 voter	1 voter
A	D	D	B	C
B	B	C	D	D
C	C	A	A	B
D	A	B	C	A

(a) Find the Condorcet winner.

(b) Find the winner according to the Coombs system.

13. **Insincere voting under Borda count.** Consider three candidates running for office, $A-C$, and 20 voters with the following preferences over candidates. Assume that the winner is selected using Borda count.

6 voters	5 voters	5 voters	4 voters
A	A	B	C
B	C	C	B
C	B	A	A

(a) Find the winner of the election if all voters assign 3 points to her most preferred candidate, 2 points to the second most preferred candidate, and 1 point to her least preferred candidate. We refer to this voting strategy as "sincere voting".

(b) How can the 4 voters in the last column alter the outcome of the election by misreporting their true preferences over candidates, i.e., insincere voting? An example suffices.

14. **Copeland's voting method.** Consider that candidates are confronted pairwise using majority voting. If a candidate i beats another candidate j according to pairwise majority rule,

he receive one point, if he ties with candidate j he receives half point. After comparing candidates in all possible pairs, we add up the number of points each candidate earned, and the winner is the candidate with the most points.

Consider the following preference profile.

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
A	E	A	D	B
C	C	B	E	D
B	A	C	C	C
D	B	D	A	E
E	D	E	B	A

(a) Find the winner according to the Copeland method.

(b) Find the winner according to the Borda count.

(c) Find the Condorcet winner.

(d) Find the winner according to majority rule.

15. **Cyclicality in bidimensional ranking.** Consider that the space of alternatives is bidimensional and, in particular, given by the unit square, i.e., $X = [0,1]^2$. A specific alternative is, hence, represented now by a pair $x = (x_1, x_2)$, rather than a point in the real line. In this setting, consider three individuals with the following utility functions:

$$\begin{aligned} u_1(x_1, x_2) &= -2x_1 - x_2, \\ u_2(x_1, x_2) &= x_1 + 2x_2, \text{ and} \\ u_3(x_1, x_2) &= x_1 - x_2. \end{aligned}$$

(a) Find the indifference curve of every individual i for a given utility level $\overline{u}$. Are his preferences convex?

(b) Show that no Condorcet winner exists. That is, demonstrate that, starting from any pair $x = (x_1, x_2)$ you can find another pair $y = (y_1, y_2)$ which is preferred by at least two of the three individuals. Importantly, you must show this result for all possible positions of pair $x = (x_1, x_2)$ on the unit square.

16. **Social welfare functions.** Consider an economy with two individuals, 1 and 2. Every individual i's utility function is $u^i(x) = \alpha^i x^i$, where $\alpha^i > 0$, and x^i represents individual i's wealth, where $x^1 + x^2 = x$.

(a) Find the socially optimal wealth distribution, i.e., the pair of wealth levels (x^1, x^2) that maximizes the social welfare function

$$W(u^1, u^2) = \left(u^1\right)^\theta + \left(u^2\right)^\theta \quad \text{where } \theta \in (0,1)$$

(b) *Numerical example.* Use the social welfare function of part (a), but assume that $\alpha^1 = 1$, $\alpha^2 = \alpha$ where $\alpha \in (0, 1)$, and $\theta = 1/3$. Identify the socially optimal wealth levels x^1 and x^2.

17. **Rawlsian swf - Properties.** Consider a Rawlsian social welfare function $W(x) = \min \left\{ u^1(x), ..., u^N(x) \right\}$.

    (a) Show that $W(x)$ is utility-level invariant.

    (b) Show that $W(x)$ satisfies anonymity.

18. **Rawlsian social welfare function - More properties.** Consider a strictly increasing and continuous social welfare function W.

    (a) Show that if W satisfies Hammond equity it can only can be represented with the Rawlsian form, $W(x) = \min \left\{ u^1(x), ..., u^N(x) \right\}$.

    (b) Show that if W is represented by the Rawlsian form, $W(x) = \min \left\{ u^1(x), ..., u^N(x) \right\}$, then W must satisfy Hammond equity.

19. **Rawlsian swf satisfying UDI.** Consider a society evaluating two alternatives x and y according to a Rawlsian swf. In particular, assume that $u^1(x) = 6$ and $u^2(x) = 12$ for alternative x, and $u^1(y) = 4$ and $u^2(y) = 12$ for alternative y.

    (a) Find the alternative that yields the highest social welfare.

    (b) Let us now apply a linear, but potentially asymmetric, strictly increasing transformation $\psi^i \left(u^i(x) \right) = a^i + bu^i(x)$, where $b = 1$. Identify for which values of parameters a^1 and a^2 utility-difference invariance (UDI) holds, and for which values this property does not hold.

20. **Utilitarian social welfare function - Properties.** Consider a society with N individuals, with a social welfare function W which is strictly increasing and continuous in every individual i's utility level u_i.

    (a) Show that if the social welfare function is utilitarian, $W(x) = \sum_{i=1}^{N} u^i(x)$, it satisfies anonymity and utility-difference invariance.

    (b) Now show the opposite line of implication. That is, show that if the social welfare function satisfies anonymity and utility-difference invariance, then W can only be represented with the utilitarian form.

21. **Flexible-form social welfare function.** In the analysis of certain policies, e.g., moving from alternative x to y, we might be interested in the percentage change in utility that each

individual experiences, $\frac{u^i(x)-u^i(y)}{u^i(x)}$, and whether such a percentage is larger for individual i than for j. That is,

$$\frac{u^i(x)-u^i(y)}{u^i(x)} > \frac{u^j(x)-u^j(y)}{u^j(x)}$$

If we seek to maintain the ranking of percentage changes across individuals invariant to monotonic transformations on the utility functions we need monotonic transformations to be *linear* and *common* among individuals, $\psi(u^i) = bu^i$, where $b > 0$ for all i. Indeed, applying this transformation on the above inequality yields

$$\frac{bu^i(x)-bu^i(y)}{bu^i(x)} > \frac{bu^j(x)-bu^j(y)}{bu^j(x)}$$

which, factoring b out, reduces to

$$\frac{u^i(x)-u^i(y)}{u^i(x)} > \frac{u^j(x)-u^j(y)}{u^j(x)}$$

Formally, we say that a social welfare function is "utility-percentage invariant" if it is invariant to arbitrary, but linear and common, strictly increasing transformations of the form $\psi(u^i) = bu^i$, where $b > 0$ for every individual i.

(a) Show that if a social welfare function satisfies utility-percentage invariant (UPI), it must also satisfy ULI and UDI. A verbal discussion suffices.

(b) Demonstrate that a strictly increasing social welfare function satisfying UPI must yields homothetic social indifference curves.

22. **Checking properties on a social welfare function - Kaneko and Nakamura (1979)**[29]
Consider the following social welfare function

$$SW(x) = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot ... \cdot x_N^{\alpha_N} = \prod_{i=1}^{N} x_i^{\alpha_i}$$

where $x = (x_1, x_2, ...x_N)$ denotes an alternative, where $x \in \mathbb{R}^N$, and $\alpha_i > 0$ represents the weight that the social planner assigns to agent i. Alternatively, this social welfare function can be represented in its linear form by applying logs, as follows,

$$\alpha_1 \ln x_1 + \alpha_2 \ln x_2 \cdot ... \cdot \alpha_N \ln x_N = \sum_{i=1}^{N} \alpha_i \ln x_i$$

In this exercise, we show that ordinal preferences among the alternatives can be represented by the above (cardinal) social welfare function that satisfies:

- Pareto Optimality (P),

[29] Kaneko M. and Nakamura K. (1979). The Nash Social Welfare Function. *Econometrica, 47*(2), 423-35.

- Anonymity (A),
- Neutrality (N),
- Independence of Irrelevant Alternatives (IIA), and
- Convexity (C).

Show that the above social welfare function $SW(x)$ satisfies these five properties.

23. **A benevolent philantropist distributing wealth.** Consider a benevolent philanthropist allocating his wealth M among his three successors. Each successor, denoted by the subscript i, has a non-zero endowment ω_i to begin with, and is bestowed for an amount of m_i from the philanthropist. The philanthropist, subject to the resource constraint $m_1 + m_2 + m_3 = M$, has assigns a weight α_i to successor i. Without loss of generality, we normalize the preference weights to $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Also, for simplicity, we assume that each successor i has a logarithmic utility function, that is, $u(x_i) = \ln x_i$.

    (a) Solve the wealth allocation problem of the philanthropist.

    (b) What are the final wealth of the successors? What is the effect of the philanthropist's preference towards successor i on his final wealth distribution? What happen if the philanthropist is "fair" in a sense of having the same preferences on all successors?

    (c) *Numerical Example.* Assume parameter values $\alpha_1 = \frac{1}{4}$, $\alpha_2 = \frac{7}{20}$, $\alpha_3 = \frac{2}{5}$, $\omega_1 = 100$, $\omega_2 = 200$, $\omega_3 = 300$, and $M = 500$. Find the wealth allocation to the successors and their final wealth respectively. What if the philanthropist is "fair"?

24. **Gibbard-Satterthwaite theorem.** In this chapter, we analyzed the aggregation of individual preferences into a social preference relation satisfying a set of desirable properties. However, we assumed individual preferences were truthfully reported by each individual. In this exercise, we examine a setting in which individuals do not necessarily truthfully reveal their preferences. In particular, we are interested in social choice functions that are "strategy proof." First, note that a *social choice function* $c\left(\succsim^1, \succsim^2, ..., \succsim^N\right) \in X$ maps the profile of individual preferences $\left(\succsim^1, \succsim^2, ..., \succsim^N\right)$ into an alternative $x \in X$. That is, society uses the social choice function (scf) to "select" an alternative $x \in X$, using the information in the profile of individual preferences $\left(\succsim^1, \succsim^2, ..., \succsim^N\right)$. Hence, we say that a scf $c(\cdot)$ is *strategy-proof* if every individual i prefers the alternative that the scf selects when he reports his true preferences, $c\left(\succsim^i, \succsim^{-i}\right) = x$, than that arising when he misreports his preferences, $c\left(\succsim'^i, \succsim^{-i}\right) = y$, i.e., $x \succsim^i y$, where $\succsim^{-i}$ denotes the profile of individual preferences by all other agents $\left(\succsim^1, ..., \succsim^{i-1}, \succsim^{i+1}, ..., \succsim^N\right)$. In words, if a scf is strategy proof, individuals have no strict incentives to misreport their preferences, regardless of the preferences other individuals report, $\succsim^{-i}$; which holds true even if the other individuals misreport their preferences. We seek to show, in several steps, Gibbard-Satterthwaite's theorem, which says that: If there

are three or more alternatives in X, then every strategy-proof scf is dictatorial.[30] In the next questions of this exercise, we will start showing that (1) a strategy-proof scf must exhibit two properties: Pareto efficiency and monotonicity; and (2) every Pareto efficient and monotonic scf must be dictatorial.

We of course need to define what we mean by Pareto efficient scf: A scf is *Pareto efficient* when every individual i's strict preference for x over y, $x \succ^i y$, where $x, y \in X$, yields the scf to select x, i.e., $c\left(\succsim^1, \succsim^2, ..., \succsim^N\right) = x$. We also define what we mean by monotonic scfs: Consider a initial profile of individual preferences, $\left(\succsim^1, \succsim^2, ..., \succsim^N\right)$, yielding that alternative x is chosen by the scf, i.e., $c\left(\succsim^1, \succsim^2, ..., \succsim^N\right) = x$. Assume that the preferences of at least individual i change from $x \succsim^i y$ to $x \succ'^i y$, for every $y \in X$, i.e., alternative x rises to the only spot at the top of his ranking of alternatives, and the preference for x is not lowered for any individual, i.e., $x \not\prec y$. We then say that a scf is *monotonic* if the scf still selects x under the new profile of individual preferences, $c\left(\succ'^1, \succ'^2, ..., \succ'^N\right) = x$. Hence, loosely speaking, a scf is monotonic if it keeps selecting x as socially preferred when x becomes the top alternative for at least one individual.

(a) Show that strategy-proofness implies monotonicity on the scf.

(b) Use monotonicity to show that the scf must be Pareto efficient.

(c) *Step 1.* Consider a profile of strict rankings in which alternative x is ranked highest and y lowest for every individual i; as illustrated in the next table. In this setting, Pareto efficiency implies that the scf must select x.

$\succsim^1$	...	$\succsim^{n-1}$	$\succsim^n$	$\succsim^{n+1}$	...	$\succsim^N$	Social choice
x	...	x	x	x	...	x	x
.		.	.	.		.	
.		.	.	.		.	
.		.	.	.		.	
y	...	y	y	y	...	y	

Consider now that we change individual 1's ranking by raising y in it one position at a time. Show that there must exist an individual n for which the social ranking changes when y is raised above x in individual n's ranking.

(d) *Step 2.* Consider now a different profile of individual preferences in which: x is moved to the bottom of individual i's ranking, for all $i < n$, and x is moved to the second last position in individual i's ranking, for all $i > n$. Show that this change in individual preferences does not change the selection of the scf.

[30] The definition of a dictatorial scf is similar to , in the definition in swf. In particular, we say that a scf $c(\cdot)$ is *dictatorial* if there is an individual d (the dictator) such that, if $x \succsim^d y$ for every two alternatives $x, y \in X$, then the scf selects x, i.e., $c\left(\succsim^1, \succsim^2, ..., \succsim^N\right) = x$. That is, a scf is dictatorial if there is an individual d such that $c(\cdot)$ chooses d's top choices, regardless of the preferences of all other individuals.

(e) *Step 3.* In this step, we use the assumption that the number of elements in the set of alternatives X is equal or larger than 3. For that, we only need to consider an alternative $z \neq x, y$ in our above steps.

(f) *Step 4.* Consider a profile of individual preferences compatible with those in Step 3. Switch the ranking of alternatives x and y for all individuals $i > n$; as depicted in the next table.

$\succsim^1$	...	$\succsim^{n-1}$	$\succsim^n$	$\succsim^{n+1}$	...	$\succsim^N$	Social choice
.		.	x	.	...	.	x
.		.	z	.		.	
.		.	y	.		.	
z	...	z		z	...	z	
y	...	y	.	y	...	y	
x	...	x	.	x	...	x	

Show that alternative x must be socially selected.

(g) *Step 5.* Argue that the scf must be dictatorial.