# Categoric and Ordinal Voting: An Overview

Harrie de Swart[1], Ad van Deemen[2], Eliora van der Hout[1] and Peter Kop[3]*

[1] Tilburg University, Faculty of Philosophy, P.O. Box 90153
5000 LE Tilburg, The Netherlands; e-mail: H.C.M.deSwart@uvt.nl
http://www.uvt.nl/faculteiten/fww/medewerkers/swart
[2] Nijmegen School of Management, University of Nijmegen, The Netherlands
e-mail: A.vanDeemen@nsm.kun.nl
[3] Faculty of Mathematics, University of Leiden, The Netherlands

**Abstract.** There are many ways to aggregate individual preferences to a collective preference or outcome. The outcome is strongly dependent on the aggregation procedure (election mechanism), rather than on the individual preferences. The Dutch election procedure is based on proportional representation, one nation-wide district, categoric voting and the Plurality ranking rule, while the British procedure is based on non-proportional representation, many districts, categoric voting and the Plurality choice rule to elect one candidate for every district. For both election mechanisms we indicate a number of paradoxes. The German hybrid system is a combination of the Dutch and British system and hence inherits the paradoxes of both systems. The STV system, used in Ireland and Malta, is based on proportional representation (per district) and on ordinal voting. Although designed with the best intentions - no vote should be wasted - , it is prone to all kinds of paradoxes. May be the worst one is that more votes for a candidate may cause him to lose his seat. The AV system, used in Australia, is based on non-proportional representation (per district) and on ordinal voting. It has all the unpleasant properties of the STV system. The same holds for the French majority-plurality rule. Arrow's impossibility theorem is presented, roughly saying that no 'perfect' election procedure exists. More precisely, it gives a characterization of the dictatorial rule: it is the only preference rule that is IIA and satisfies the Pareto condition. Finally we mention characterizations of the Borda rule, the Plurality ranking rule, the British FPTP system and of $k$-vote rules.

## 1 Introduction

In this overview, we give an analysis of election procedures and their properties. An election mechanism can serve, given individual preference orderings of the

* We thank Marc Roubens for some quite useful suggestions and Sven Storms for his help in translating the original Dutch manuscript. This paper is an extended version of the original Dutch booklet 'Verkiezingen', published in 2000 by Epsilon Uitgaven, Utrecht, The Netherlands. We thank Epsilon Uitgaven for permission to do so.

alternatives, to select one alternative, for instance, a travel goal or a chairman. In these cases, we speak of a (collective) *choice rule*. An election procedure can also be used to select a set of alternatives, for instance, a parliament or a set of potential bus stops. In this case, we speak of a (collective) *choice correspondence*. Finally, an election mechanism can be used to determine an order of collective preference regarding the alternatives, for instance, of candidates for the Eurovision Song Contest. In this case, we speak of a (collective) *preference rule*.

In section 2, it becomes clear that the outcome of elections is strongly dependent on the election procedure used. We consecutively consider: Most votes count (Plurality Rule), Pairwise comparison (Majority Rule), the Borda rule, and Approval voting. There are numerous other election procedures, too many to name here.

In sections 3 and 4, we distinguish four different kinds of election procedures that are used in most Western European countries to elect parliament and government. Subsequently, we show that each of the four globally distinguished election procedures is subject to paradoxes. By 'paradox' we mean an outcome that is contrary to what one would prima facie expect or contrary to our sense of justice and honesty. For instance, it is a paradox that more votes for a candidate or party under a specific election procedure can mean that the candidate or party is worse off. (This is the Negative Responsiveness paradox for the election procedure designated by STV.)

In Section 3, we compare the Dutch election procedure to the British one. In section 3.2, paradoxes are discussed that occur, or can occur, within the Dutch system. Section 3.3 considers the paradoxical properties of the British system. The hybrid election procedure that is used in Germany is treated in section 3.4. This system, which is a combination of the Dutch and the English systems, also has its own paradoxes.

The Single Transferable Vote (STV) and the Alternative Vote (AV) systems are discussed in Section 4. Section 4.2 elaborates on the properties of the STV election procedure that is used in Ireland and Malta. Section 4.4 considers some paradoxes that may occur in the election procedure that is used in Australia. Finally, in section 4.5, we deal with the French election system, which is very similar to the AV system that is used in Australia and is similarly the cause of several paradoxes.

Naturally, the question then arises if there are any 'good' election procedures, that is, election procedures that, at any rate, do not have the unwanted properties that we noted in the chapters mentioned above. Kenneth Arrow addressed this question over fifty years ago. In Section 5, we examine Arrow's result, which is essentially a characterization of the dictatorial rule. Although no 'perfect' election procedure exists, some procedures are 'better' than others. One way to decide on this, is by studying the characteristic properties of these procedures. We mention characterizations of the Plurality ranking rule, of the Borda rule and of $k$-vote rules.

## 2 Other Procedure, Other Outcome

In this section, we consider a number of election procedures. These are procedures by which the outcome of an election is determined. At first glance, you might think that this is simple: most votes count. Doesn't that seem fairest? However, we will see that there are objections to the 'Most votes count' (Plurality Rule) election procedure. Hence, we also look at other procedures: the Majority Rule, the Borda rule, and Approval Voting.

For all our examples, we assume that we know the individual preferences of the voters. A survey (the whole) of all individual preferences is called a *(voter)profile*, denoted by the symbol $p$ or $q$. An election procedure is a procedure that assigns to each (voter)profile an outcome (of the election).

In Example 1 (see below), we will show that different election procedures may produce different outcomes. This means that one can doubt whether the outcome generated by any single procedure is the 'best' or 'correct' outcome. In other words, one can doubt the appropriateness and quality of the used procedure.

**Example 1**: A group of secondary school students is given the choice between Venice, Florence, and Siena as the destination of their school trip. Each student is allowed to give his or her order of preference, for instance,

Venice Siena Florence.

This means that Venice is the first preference of this student, Siena the second, and Florence the third. Now suppose that there are 31 students with the following individual preferences.

| | | | |
|---|---|---|---|
| Florence | Venice | Siena | : 5 students |
| Florence | Siena | Venice | : 7 students |
| Venice | Florence | Siena | : 3 students |
| Venice | Siena | Florence | : 7 students |
| Siena | Florence | Venice | : 3 students |
| Siena | Venice | Florence | : 6 students |

Such a survey of individual preferences is called a *profile*, usually denoted by the letter $p$. Election procedures aggregate profiles of individual preferences to an outcome.

In this example, if each student is allowed to give his or her first preference and the procedure 'Most votes count' is applied, then Florence, with $5 + 7 = 12$ votes, will be selected. Later, we will see that other election procedures might assign different outcomes to this same (voter)profile.

We also discuss a number of important properties of election procedures in this chapter, such as the Pareto condition, the condition of Independence of Irrelevant Alternatives (IIA), and the monotonicity-condition. We will explain these conditions using examples.

### 2.1 Plurality Rule

The election procedure ‘Most votes count’ only considers the first preferences of the voters; second, third, etc., preferences are not considered. For ‘Most votes count’ (Plurality Rule), alternative $x$ is *collectively* (by the community) *preferred* to alternative $y$ if the number of persons that prefer $x$ is greater than the number of persons that prefer $y$. In particular, if one choice is needed, the alternative that is put first by most people will be elected. We call $x$ and $y$ (collectively) *indifferent* if the number of individuals that prefer $x$ is equal to the number of individuals that prefer $y$.

If there are just two alternatives, or candidates $x$ and $y$, ‘$x$ is collectively preferred over $y$’ means that $x$ gets more than half of the (first) votes.

In the (voter)profile of Example 1, Florence is mentioned 12 times as first preference, Venice 10 times, and Siena 9 times. Therefore, on application of the ‘Most votes count’ election procedure, Florence will become - as we already saw - the destination of our class. In other words, Florence is the (collective) choice of our class under application of the ‘Most votes count’ election procedure.
Not only can ‘Most votes count’ be used to determine a collective choice, but also to determine a collective order of preferences. In that case one speaks of the *Plurality ranking rule*. Given the profile of Example 1, the collective order of preference on application of ‘Most votes count’ will be

Florence Venice Siena.

This corresponds to the fact that Florence gets more first votes than Venice, which in turn gets more first votes than Siena.

Suppose that later on it turns out that Venice is so expensive that it was not a realistic alternative. One could then argue that a new vote is not needed, as Venice was not the chosen destination anyway. However, if Venice is no longer an alternative and the preferences of the students remain unchanged as far as the other alternatives are concerned, the preferences of the 31 students will be as follows:

| | | |
|---|---|---|
| Florence | Siena | : 5 students |
| Florence | Siena | : 7 students |
| Florence | Siena | : 3 students |
| Siena | Florence | : 7 students |
| Siena | Florence | : 3 students |
| Siena | Florence | : 6 students |

Now there are 15 students with Florence as first preference and 16 with Siena as first preference. So, on application of ‘Most votes count’, Siena would be elected as the destination instead of Florence.

We say that ‘Most votes count’ is not *Independent of Irrelevant Alternatives* (not IIA): although Venice is an irrelevant alternative, because of the cost, the outcome is not independent of this alternative. The property ‘Independent of

Irrelevant Alternatives (IIA)' can also be described as follows: adding irrelevant (non eligible) alternatives does not influence the outcome.

'Most votes count' is frequently used in real life: it is the foundation of many election systems that are in current use, such as the Dutch and British systems (see section 3.1). Nonetheless, this system has some serious drawbacks, as we will explain below and in subsections 3.2 and 3.3. (Here, we will follow the exposition of Van Deemen, 1997).

In the first place, a choice made using the procedure 'Most votes count' is not necessarily a majority choice. This remarkable fact was discovered as early as 1781 by the Frenchman J.-C. de Borda (1781), one of the founders of Social Choice Theory. To clarify this, we consider the voter profile of Example 1.

- Check that there are $0 + 0 + 3 + 7 + 0 + 6 = 16$ students that prefer Venice to Florence, and 15 that prefer Florence to Venice. In other words, if the students have to choose between Florence and Venice, they will choose Venice.
- Check that there are $0 + 0 + 0 + 7 + 3 + 6 = 16$ students that prefer Siena to Florence, and 15 that prefer Florence to Siena.
- Also check that there are $0 + 7 + 0 + 0 + 3 + 6 = 16$ students that prefer Siena to Venice, and 15 that prefer Venice to Siena.

Hence, we can conclude that
1) On pairwise comparison, Florence has a minority of the votes with respect to both Venice and Siena: for this reason, Florence is called a *Condorcet loser*, after the French Marquis de Condorcet (1743 - 1794).
2) On pairwise comparison, Siena has a majority of the votes with respect to both Florence and Venice, and, hence, Siena is the majority choice of our class; for that reason, Siena is called the *Condorcet winner*.

From the above, it follows that the winner on application of 'Most votes count' (Florence) need not be the majority choice (Siena). In other words, the *majority principle* is violated by 'Most votes count'.

To clarify the second drawback of 'Most votes count', we consider the following voter profile.

| | | | | | |
|---|---|---|---|---|---|
| Florence | Paris | London | Venice | Siena | : 10 voters |
| Siena | Paris | Venice | Florence | London | : 8 voters |
| Venice | Siena | Paris | London | Florence | : 7 voters |

As neither Paris nor London is the first preference of any voter, they are collectively indifferent on application of 'Most votes count': for each city, the number of individuals for whom it is first choice is 0. However, everyone prefers Paris to London. How odd! Everyone prefers Paris to London, but this is not shown in the outcome: Paris and London are equally preferred in the outcome.

The aforementioned comes down to the fact that the election procedure 'Most votes count' violates the so-called Pareto condition. This *Pareto condition* goes as follows: if every individual prefers alternative $x$ to alternative $y$, then, in the outcome, $x$ must also be (collectively) preferred to $y$.

The third drawback of 'Most votes count' is that it does not have the monotony property. This *monotony property* (positive responsiveness) says that if an alternative $x$ is raised vis-a-vis an alternative $y$ in someone's preference ordering, and $x$ goes down in no one's preference ordering vis-a-vis $y$, then $x$ must also be raised vis-a-vis $y$ in the collective preference ordering. To see that 'Most votes count' does not have this monotony property, we consider the following (voter)profile $p$. With $(xy)$ we mean that $x$ and $y$ are indifferent in the preference ordering.

| | | | | | |
|---|---|---|---|---|---|
| Profile $p$: | Florence | (Paris London) | Venice | Siena | : 10 students |
| | Siena | (Paris London) | Venice | Florence | : 8 students |
| | Venice | (Paris London) | Siena | Florence | : 7 students |

Because neither Paris nor London occurs as first preference in the preference ordering of the students, on application of 'Most votes count' both are indifferent. But now consider the following profile $q$, identical to profile $p$ except for the fact that everybody now prefers Paris to London in his or her preference ordering.

| | | | | | | |
|---|---|---|---|---|---|---|
| Profile $q$: | Florence | Paris | London | Venice | Siena | : 10 students |
| | Siena | Paris | London | Venice | Florence | : 8 students |
| | Venice | Paris | London | Siena | Florence | : 7 students |

Comparing the profiles $p$ and $q$, in profile $q$ everybody has ranked Paris higher in his or her preference ordering than London. So, according to the monotony property, Paris should now be (collectively) preferred by the community to London. However, on application of 'Most votes count', this is not the case, since neither Paris nor London is the first preference of an individual and, hence, they are indifferent in the collective preference (if this is determined by 'Most votes count'). Consequently, the election procedure 'Most votes count' may not react to changes in the individual preferences, which seems at odds with the idea of democracy.

Given these results, it is no wonder that Borda and Condorcet had little faith in 'Most votes count'!

## 2.2 Profiles, choice, and preference rules

In this subsection, we will formulate in a mathematically precise way a number of properties that were introduced informally in the previous subsection, as well as add some new mathematical notions. Amongst others, the following concepts will be defined: relation, weak and linear ordering, profile, choice rule, choice correspondence, preference rule, and Independence of Irrelevant Alternatives.

The individual order of preference 'Florence Venice Siena' can be rendered by the following *(preference-)relation* $R$:

$$R = \{<\text{Florence, Venice}>, <\text{Venice, Siena}>, <\text{Florence, Siena}>\}.$$

Here $< x, y >$ is an *ordered pair* , and $< x, y > \in R$ is read as '$x$ is *at least as good* as $y$'. Instead of $< x, y > \in R$, we usually write $xRy$.
'$x$ is (strictly) *preferred to* $y$' now corresponds with '$xRy$ and not $yRx$', while '$xRy$ en $yRx$' states that '$x$ and $y$ are *indifferent*', which is often denoted by $(xy)$.

Suppose that $A$ is a set of alternatives, for instance, $A = \{$Florence, Venice, Siena$\}$ and that $N$ is a set of individuals, for instance, $N = \{$student 1, ..., student 31$\}$. Then we can identify for every individual $i$ in $N$ his or her individual preference ordering with respect to the alternatives in $A$ by means of a *relation* $R_i$ on $A$, also called a *preference-relation* on $A$.

**Definition 1** $R$ is a *(preference-)relation* on $A$ if $R$ is a set of ordered pairs $< x, y >$ with $x, y \in A$. Instead of writing $< x, y > \in R$, one can also write $xRy$.

**Definition 2** Let $R$ be a (preference-)relation on $A$.
$R$ is *complete* if, for all $x, y \in A$, $xRy$ or $yRx$. That is, a relation on $A$ is complete if every alternative in $A$ is comparable to every alternative in $A$, including itself. Recall that $xRy$ is read as '$x$ is at least as good as $y$'.
$R$ is *transitive* if, for every $x, y, z \in A$, if $xRy$ and $yRz$, then $xRz$. That is, if $x$ is at least as good as $y$ by $R$ and $y$ is at least as good as $z$ by $R$, then $x$ is at least as good as $z$ by $R$. Thus, in transitive preference relations, the preferences are consequent.
$R$ is *antisymmetric* if, for every $x, y \in A$ with $x \neq y$, if $xRy$, then not $yRx$. That is, a relation is antisymmetric if indifference between two distinct alternatives does not occur. '$xRy$ and not $yRx$' is read as: $x$ is (strictly) preferred to $y$ by $R$.

**Definition 3** A preference relation $R$ is a *weak ordering* on $A$ if $R$ is complete and transitive. $R$ is a *linear ordering* on $A$ if $R$ is complete, transitive, and antisymmetric. Hence, there can be indifference in weak orderings, but not in linear orderings.

**Definition 4** $C(A)$ is, by definition, the set of all complete relations on $A$. $W(A)$ is, by definition, the set of all weak orderings on $A$. $L(A)$ is, by definition, the set of all linear orderings on $A$. Because every linear ordering is, by definition, also a weak ordering, it follows that $L(A)$ is a subset of $W(A)$, while $W(A)$, in its turn, is a subset of $C(A)$: $L(A) \subseteq W(A)$ and $W(A) \subseteq C(A)$.

For the sake of simplicity, we will limit ourselves to individual preferences $R_i$ that are linear orderings. With a *profile* $p$, we mean a combination of individual linear orderings.

**Definition 5** A *profile* $p$ associates with every individual $i$ in $N$ a linear ordering $R_i$ on $A$, in other words,

$$\text{a } \textit{profile} \text{ is a function } p : N \to L(A).$$

$p(i)$ or $R_i$ is the individual linear ordering of individual $i$ in profile $p$.
$L(A)^N$ is the set of all profiles.

So, in Example 1, a profile is given for which
$N = \{\text{student 1,..., student 31}\}$ and $A = \{\text{Florence, Venice, Siena}\}$.

A group of individuals can make three kinds of collective decision on the basis of a given voter profile (a combination of individual preferences).

1. It can choose one alternative, for instance, a travel destination, a chairman, a president, or a location for a sporting facility.
2. It can choose a collection of alternatives, for instance, a parliament, a food package, or a set of potential locations for a waste dump.
3. It can determine an order of preference of the alternatives, for instance, of applicants or of candidates for the Eurovision Song Festival.

In Case (1), we call the election procedure a *(collective) choice rule*, in Case (2), we call it a *(collective) choice correspondence*, and, in Case (3), we call it a *(collective) preference rule.*

**Definition 6** Let $N$ be a set of individuals and $A$ a set of (at least 3) alternatives.

1. A *(collective) choice rule* is a function $K : L(A)^N \to A$. Thus, a choice rule $K$ assigns to each profile $p \in L(A)^N$ a collective choice $K(p)$ in $A$.
2. A *(collective) choice correspondence* is a function $C : L(A)^N \to P(A)$, where $P(A)$ is the powerset of $A$. This is the collection of all subsets of $A$. Therefore, a choice correspondence $C$ assigns to each profile $p \in L(A)^N$ a set $C(p)$ of collective choices in $A$.
3. A *(collective) preference rule* is a function $F : L(A)^N \to C(A)$. Thus, a preference rule $F$ assigns to each profile $p \in L(A)^N$ a complete preference relation $F(p)$ on $A$.

The election procedure 'Most votes count' can be seen as a (collective) choice rule or choice correspondence and as a (collective) preference rule.

**Definition 7** Suppose $N$ is a set of individuals and $A$ is a set of alternatives. Given a profile $p$ and an alternative $x$ in $A$, we define $t(x, p)$ as the number of individuals $i$ in $N$ that have $x$ as the *first* preference in $p(i)$ (i.e., for which there is no alternative $y$ in $A$ that is more preferred by $i$ than $x$ in $p(i)$).

'Most votes count' as a preference rule is now rendered by the function $Pl$ (Plurality) from $L(A)^N$ to $W(A)$, defined as

$$xPl(p)y \text{ if and only if } t(x, p) \geq t(y, p).$$

In other words, $xPl(p)y$ if and only if the number of individuals that prefer $x$ most in $p$ is greater than or equal to the number of individuals that prefer $y$ most in $p$. Note that $Pl(p)$ is a weak ordering on $A$ and, in general, not a linear ordering, because there can be two or more alternatives that occur equally often as first preference in $p$.

**Definition 8** The collective preference rule $Pl$ gives rise to the collective choice correspondence $P$ (Plurality), $P : L(A)^N \rightarrow P(A)$, with $P(p)$, by definition, the set of all $x$ in $A$ such that, for all $y$ in $A$, $xPl(p)y$. Therefore, $P(p)$ is the set of all alternatives $x$ in $A$ for which there is no alternative $y$ in profile $p$ which is more frequently preferred most in $p$.

**Definition 9** Let $F : L(A)^N \rightarrow C(A)$ be a (collective) preference rule. $F$ is *Independent of Irrelevant Alternatives* (IIA) if, for all $x, y \in A$ and for all profiles $p, q \in L(A)^N$, if $p$ limited to $x$ and $y$ is equal to $q$ limited to $x$ and $y$, then $F(p)$ limited to $x$ and $y$ is equal to $F(q)$ limited to $x$ and $y$.

So, if $p$ is the profile from Example 1 and $q$ is the same profile but without Venice or with Venice as last choice, then $p$ limited to Florence and Siena is equal to $q$ limited to Florence and Siena. Now, let $Pl$ (Plurality) be the (collective) preference rule that corresponds to 'Most votes count'. $Pl(p) = \{$<Florence, Venice>, <Venice, Siena>, <Florence, Siena>$\}$ and $Pl(q) = \{$<Siena, Florence>, <Siena, Venice>, <Florence, Venice>$\}$. Then, $Pl(p)$ limited to Florence and Siena would be $\{$<Florence, Siena>$\}$, but $Pl(q)$ limited to Florence and Siena would be $\{$<Siena, Florence>$\}$. So, $Pl$, which is 'Most votes count', is not Independent of Irrelevant Alternatives (not IIA).

### 2.3 Majority Rule (pairwise comparison)

The *majority principle* states that if the number of voters that prefer alternative $x$ to alternative $y$ is larger than the number of voters that prefer $y$ to $x$ (in other words, if $x$ *defeats* $y$), then $x$ must also be preferred to $y$ in the outcome. It follows from this that, if there is an alternative $x$ that defeats every other alternative in pairwise comparison, this alternative $x$ must win. Such an alternative is called a *Condorcet winner* .

In the previous section, we saw how, given a voter profile, the Condorcet winner is determined and that this Condorcet winner need not be the winner under application of 'Most votes count'. In fact, with the voter profile of Example 1, the winner under application of 'Most votes count' (Florence) is the *Condorcet loser*: on pairwise comparison, Florence loses from both Venice and Siena.

It is difficult to justify the fact that a candidate or party preferred by a minority, may get elected or receive more seats than a candidate or party that is preferred by a majority. Therefore, Borda (1781) and Condorcet (1788) concluded that the procedure 'Most votes count' is seriously defective, because it does not satisfy the majority principle.

As the majority principle seems so plausible, one could wonder why we still use other procedures. The answer is simple: there are profiles that have no Condorcet winner. The most famous example is the following so called *Condorcet profile* $p$ (in which $k$ is a random natural number, $k \geq 1$):

| | | | |
|---|---|---|---|
| Florence | Venice | Siena | : $k$ students |
| Venice | Siena | Florence | : $k$ students |
| Siena | Florence | Venice | : $k$ students |

The Majority Rule (pairwise comparison) applied to the above Condorcet profile leads to a collective order of preference that is *not transitive*, meaning that alternatives $x$, $y$, and $z$ exist, in our example, respectively, Florence, Venice, and Siena, such that $x$ defeats $y$ and $y$ defeats $z$, but $x$ does not defeat $z$. Hence, for the above Condorcet profile, no Condorcet winner can be found.

The absence of a Condorcet winner for a profile is also called the *Condorcet paradox* or *voting paradox*.

One might wonder if the probability of an occurrence of the Condorcet paradox in actual elections is significantly large. Bill Gehrlein (1981) showed that, under certain assumptions, the probability, in the case of three alternatives, is $\frac{1}{16}$ if the number of individuals is large. For more than three alternatives, the probability of the Condorcet paradox occurring increases; see [16].

Despite the Condorcet paradox, the Majority Rule (pairwise comparison) has a number of properties that come close to the ideal of a democracy. In [10], 3.2.1, Van Deemen notes that the Majority Rule (pairwise comparison) has the following properties.

– *Anonymity*: Individuals are treated equally. It does not matter from whom the preferences originated, the only thing that counts are the preferences themselves. Personal qualifications of the individuals are irrelevant to the determination of the collective choice. Anonymity prevents unequal treatment of individuals: it erects a barrier to any form of discrimination. Note that 'Most votes count' is also an anonymous election procedure.
– *Neutrality*: The alternatives are treated equally. Every opinion counts, independent of its content. Note that 'Most votes count' also has this property.
– *Independence of Irrelevant Alternatives (IIA)*: The determination of the collective preference with respect to two alternatives $x$ and $y$ is not influenced by a third (irrelevant) alternative. In Section 2.1, we have seen that 'Most votes count' is not IIA.
– *Pareto condition*: If everybody prefers alternative $x$ to alternative $y$, then $x$ will also be collectively preferred to $y$. In Section 2.1, we have seen that 'Most votes count' does not satisfy the Pareto condition.
– *Monotony*: If an alternative $x$ is raised vis-a-vis an alternative $y$ in someone's preference ordering and $x$ goes down in no one's preference vis-a-vis $y$, then, on pairwise comparison, $x$ will also be raised vis-a-vis $y$ in the collective order of preference. A voting procedure that does not have this property can be regarded as having a certain inertia: it cannot register changes in the profiles and adapt its outcome in accordance with these changes. In Section 2.1, we showed that 'Most votes count' is not monotonic.

We can speak of an election procedure even in the case of dictatorship. An individual is called a *dictator* if, for every voter profile $p$, the collective preference is exactly the preference of that individual. The *dictatorial preference rule* with dictator $i$ assigns to each voter profile $p$ the preference of $i$. See [34], pp. 70-72.

For instance, consider a class with individual preferences as in Example 1 and a teacher with preference ordering Venice Florence Siena. If the teacher plays the role of dictator, the class will go to Venice.

Check that a dictatorial preference rule is not anonymous, but neutral, IIA, and satisfies the Pareto Condition. (In Section 5, we will see that the dictatorial preference rule is the only preference rule that is IIA and satisfies the Pareto condition.)

In the next section, we will formulate the above mentioned properties in a mathematically precise way.

### 2.4 Properties of the Majority Rule

**Definition 10** Given a profile $p$, an alternative $x$ *defeats* an alternative $y$ *on pairwise comparison* if the number of individuals that prefer $x$ to $y$ in profile $p$ is greater than the number of individuals that prefer $y$ to $x$ in $p$. Given a profile $p$, we write '$x$ defeats $y$ on pairwise comparison' as $xM(p)y$ (the $M$ stands for *Majority Rule*). This defines the collective preference rule $M : L(A)^N \to C(A)$.

A *Condorcet winner* is an alternative that defeats any other alternative on pairwise comparison.

Note that the relation $M(p)$ need not be transitive, for instance, if $p$ is a Condorcet profile. Also note that there can be several Condorcet winners. For instance, in the following profile $p$, where $(xz)$ means that $x$ and $z$ are indifferent.

$$\begin{array}{llll} z & x & y & : 3 \\ y & x & z & : 3 \\ & (xz)y & & : 1 \end{array}$$

**Definition 11** A permutation $\sigma$ of $N$ is a bijective function from $N$ to $N$. We can see a permutation $\sigma$ of $N$ as a name change for all individuals in $N$. After application of $\sigma$, individual $i$ is named $\sigma(i)$.

Let $p$ be a profile in $L(A)^N$. Then $p \circ \sigma$ is, by definition, the profile in which each individual $i$ plays the role of $\sigma(i)$ in $p$. So, for all $i$ in $N$, $(p \circ \sigma)(i) = p(\sigma(i))$.

**Example**: Suppose that $N = \{$a(d), b(ob), c(ees)$\}$ and that $\sigma$(a) = b, $\sigma$(b) = c and $\sigma$(c) = a. Suppose also that $A = \{$Florence, Venetië, Siena$\}$ and that profile $p$ is given by

| | | | |
|---|---|---|---|
| $p$(a) : | Florence | Venice | Siena |
| $p$(b) : | Florence | Siena | Venice |
| $p$(c) : | Venice | Florence | Siena |

Then, $p \circ \sigma$ is the following profile:

| | | | |
|---|---|---|---|
| $p \circ \sigma$(a) = $p$(b) : | Florence | Siena | Venice |
| $p \circ \sigma$(b) = $p$(c) : | Venice | Florence | Siena |
| $p \circ \sigma$(c) = $p$(a) : | Florence | Venice | Siena |

It can be easily seen that $M(p \circ \sigma) = M(p) = \{$<Florence, Venice>, <Venice, Siena>, <Florence, Siena>$\}$.

**Definition 12** A collective preference rule $F : L(A)^N \to C(A)$ is *anonymous* if, for all profiles $p$ in $L(A)^N$ and for every permutation $\sigma$ of $N$, $F(p \circ \sigma) = F(p)$.

**Definition 13** Suppose $\tau$ is a permutation of $A$ and $R$ is a complete relation on $A$. Then $\tau R$ is, by definition, the set of all pairs $< \tau(x), \tau(y) >$ with $< x, y >$ in $R$. So, in $\tau R$, $\tau(z)$ plays the role of $z$ in $R$.

Let $p$ be a profile in $L(A)^N$. Then, $\tau p$ is, by definition, the profile with $(\tau p)(i) = \tau(p(i))$ for all $i$ in $N$. $\tau p$ originates from $p$ by applying the permutation $\tau$ on the alternatives.

**Example**: Suppose that $N = \{$a, b, c$\}$ and $A = \{$Florence, Venice, Siena$\}$. Suppose $\tau$ is the permutation of $A$ given by $\tau$(Florence) = Venice, $\tau$(Venice) = Florence en $\tau$(Siena) = Siena. And suppose that $p$ is the following profile:

| | | | |
|---|---|---|---|
| $p$(a) : | Florence | Venice | Siena |
| $p$(b) : | Siena | Venice | Florence |
| $p$(c) : | Venice | Siena | Florence |

Then, $M(p) = \{$<Venice, Siena>, <Siena, Florence>, <Venice, Florence>$\}$.
The profile $\tau p$ now originates from profile $p$ by interchanging the alternatives Florence and Venice:

| | | | |
|---|---|---|---|
| $\tau p$(a) $= \tau(p$(a)$)$ : | Venice | Florence | Siena |
| $\tau p$(b) $= \tau(p$(b)$)$ : | Siena | Florence | Venice |
| $\tau p$(c) $= \tau(p$(c)$)$ : | Florence | Siena | Venice |

It can now be easily seen that $M(\tau p) = \tau(M(p)) = \{$<Florence, Siena>, <Siena, Venice>, <Florence, Venice>$\}$.

**Definition 14** A collective preference rule $F : L(A)^N \to C(A)$ is *neutral* if, for every permutation $\tau$ of $A$ and for every profile $p$, $F(\tau p) = \tau(F(p))$.

**Definition 15** A collective preference rule $F : L(A)^N \to C(A)$ satisfies the *Pareto condition* if, for every profile $p$ in $L(A)^N$ and for all alternatives $x, y$ in $A$, if for every $i \in N$ $xp(i)y$ (and hence not $yp(i)x$), then $xF(p)y$ and not $yF(p)x$.

**Definition 16** A collective preference rule $F : L(A)^N \to C(A)$ is *monotonic* if, for all profiles $p, q$ in $L(A)^N$ and for all alternatives $x, y$ in $A$, if
1. for all $i \in N$, if $xp(i)y$ (and hence not $yp(i)x$), then $xq(i)y$ (and hence not $yq(i)x$), and
2. there is an individual $k \in N$ such that $yp(k)x$ and $xq(k)y$,
then $xF(p)y$ implies that $xF(q)y$ and not $yF(q)x$.

As was mentioned earlier, the following theorem is easy to see.

**Theorem 1** The collective preference rule $M$ (Majority Rule) is anonymous, neutral, IIA, monotonic, and satisfies the Pareto condition, but it is not transitive.

In order to avoid the voting paradox or non-transitivity, Copeland modified the Majority Rule in the following way. The Copeland score of an alternative $x$ given profile $p$ is by definition the number of alternatives $y$ such that $x$ defeats $y$ on pairwise comparison given $p$. The Copeland preference rule $F_{Copeland}$ is now defined by $xF_{Copeland}y$ if and only if the Copeland score of $x$ given $p$ is greater than or equal to the Copeland score of $y$ given $p$. So, $x$ is more preferred than $y$ by $F_{Copeland}(p)$ if and only if $x$ defeats more alternatives than $y$ given $p$. Evidently, the Copeland preference rule is transitive, as well as anonymous and neutral, it satisfies the Pareto Condition, but it is not IIA.

### 2.5 Borda (preference) rule

In 1781, the Frenchman J.C. de Borda noted that, with 'Most votes count', the second, third, etc., preferences of the individuals have no weight in determining the outcome. Borda proposed giving weight to all the positions of the alternatives in the individual preferences. Hence, not only the first preference of the individuals is taken into account, but also their second, third, etc. If an individual $i$ has 'Florence Venice Siena' as individual preference ordering, Florence gets 3 points, Venice 2, and Siena 1. Subsequently, a decision is made based on the total score of every alternative in a given profile $p$. For $n$ alternatives, every individual gives $n$ points to his or her most preferred alternative, $n-1$ points to his or her second choice, etc., and 1 point to his or her least preferred alternative.

If we apply the Borda preference rule to Example 1 (see page 150), Florence, Venice, and Siena will get the following numbers of points:

| | |
|---|---|
| Florence: | $5 \times 3 + 7 \times 3 + 3 \times 2 + 7 \times 1 + 3 \times 2 + 6 \times 1 = 61$ |
| Venice: | $5 \times 2 + 7 \times 1 + 3 \times 3 + 7 \times 3 + 3 \times 1 + 6 \times 2 = 62$ |
| Siena: | $5 \times 1 + 7 \times 2 + 3 \times 1 + 7 \times 2 + 3 \times 3 + 6 \times 3 = 63$ |

The *Borda score* of an alternative $x$ for a given profile $p$ is now, by definition, the total number of points that the individuals have given to $x$. In Example 1, the Borda score of Florence is 61, the Borda score of Venice is 62, and the Borda score of Siena is 63.

According to the *Borda (preference) rule*, the collective ordering of the alternatives will then be

Siena Venice Florence.

Note that, for 'Most votes count', the outcome for the profile of Example 1 is exactly the opposite,

Florence Venice Siena,

because, in Example 1, Florence is preferred 12 times, Venice 10 times, and Siena 9 times. Also note that Siena, with the highest Borda score, happens to be the Condorcet winner in Example 1.

The obvious question now is if the Condorcet winner, if one exists, will always have the highest Borda score. Unfortunately, this is not the case, as is shown by the following example.

A group of seven people go out for dinner. The restaurant offers three menus: $a$, $b$, and $c$. As there is a reduction if they all take the same menu, they decide to choose collectively. But which menu should be chosen? The individual preferences are given in the profile below.

| | | | |
|---|---|---|---|
| $c$ | $a$ | $b$ | : 3 persons |
| $a$ | $b$ | $c$ | : 2 persons |
| $a$ | $c$ | $b$ | : 1 person |
| $b$ | $c$ | $a$ | : 1 person |

1) Check that $c$ is the Condorcet winner for this profile!
2) Now check that $c$, when the Borda procedure is applied to this profile, only receives 15 points, while $a$ gets 16 points under these circumstances. Thus, an alternative with the highest Borda score need not be the Condorcet winner.

The profile just mentioned also illustrates that, like 'Most votes count', the Borda procedure is not Independent of Irrelevant Alternatives (not IIA).
1) On application of the Borda procedure on the profile just given, the collective order of preference is $a\ c\ b$.
2) When they want to order menu $a$, the waiter tells them this is very convenient, as menu $b$ cannot be served today. You might think this information is unimportant. However, if the Borda procedure is applied in this new situation (only $a$ and $c$), the collective order of preference will become $c\ a$.

So, for the Borda procedure, the presence of the (irrelevant) alternative $b$ influences the preference between $a$ and $c$. Hence, the collective choice between $a$ and $c$, on application of the Borda procedure, is dependent on all alternatives, in particular on the irrelevant alternative $b$.

Note that, when there are two alternatives, the Borda procedure corresponds to 'Most votes count' as well as to the Majority Rule. Suppose that there are two alternatives, $x$ and $y$, and $m+n$ individuals, and that the individual preferences are given in the following profile:

| | | |
|---|---|---|
| $x$ | $y$ | : $m$ voters |
| $y$ | $x$ | : $n$ voters |

Then, the Borda score of $x$ equals $2m+n$ and the Borda score of $y$ equals $2n+m$. Now $2m+n > 2n+m$ if, and only if, $m > n$. Thus, the Borda score of $x$ is greater than that of $y$ precisely when the number of voters ($m$) that prefer $x$ to $y$ is greater than the number of voters ($n$) that prefer $y$ to $x$.

The reader may easily verify the following theorem.

**Theorem 2** The Borda (preference) rule is anonymous, neutral, not IIA, monotonic, transitive and satisfies the Pareto condition.

It is worth mentioning that the Majority Rule and the scoring procedure (generally ascribed to Borda) were in fact first proposed respectively by Ramon Lull (± 1235 - 1315) and Nicolas Cusanus (1401 - 1464), as reported in [26] and [28].

### 2.6 Strategic behavior

We have already seen that the Borda preference rule does not necessarily pick out the Condorcet winner, if there is one. Another drawback of the Borda rule is that it is sensitive to strategic behavior. This means that individuals can profit from giving an insincere preference instead of their true preference. To illustrate this, we consider the following profile (17 voters):

| | | | |
|---|---|---|---|
| Florence | Venice | Siena | : 7 students |
| Venice | Florence | Siena | : 6 students |
| Siena | Venice | Florence | : 4 students |

The Borda score for Siena is $7 \times 1 + 6 \times 1 + 4 \times 3 = 25$. The Borda score for Florence is $7 \times 3 + 6 \times 2 + 4 \times 1 = 37$. The Borda score for Venice is $7 \times 2 + 6 \times 3 + 4 \times 2 = 40$. So the outcome is

Venice Florence Siena.

Venice ends above Florence. The first group of 7 students, preferring Florence to Venice, can now act strategically: instead of giving their true preferences, they can vote as follows:

Florence Siena Venice

Venice now gets 7 points less: $40 - 7 = 33$, while Siena gets an extra 7 points: $25 + 7 = 32$. As the score of Florence remains unaltered, 37 points, the resulting collective ordering is now

Florence Venice Siena.

This is exactly the outcome desired by the first group of seven students.

In this example, a coalition of seven voters acts strategically and benefits from this. One could remark that the strategic behavior of a coalition presupposes internal attunement and, hence, would be difficult to realize in practice. The next example shows that one person can also benefit from strategic behavior, assuming that the other voters give their true preferences.

Suppose that there are five alternatives, $x$, $y$, $z$, $u$, and $v$, and seven voters. Also suppose that the (sincere) individual preferences are given in the following profile:

| | | | | | |
|---|---|---|---|---|---|
| $x$ | $y$ | $z$ | $u$ | $v$ | : 3 persons |
| $z$ | $x$ | $y$ | $u$ | $v$ | : 2 persons |
| $y$ | $z$ | $x$ | $u$ | $v$ | : 2 persons |

Now, the Borda score of $x$ is 29, that of $y$ 28, of $z$ 27, of $u$ 14, and of $v$ 7. So, on application of the Borda preference rule, the outcome for the above profile will be

$$x\ y\ z\ u\ v.$$

Now, suppose that one of the last two voters foresees this outcome. Now this voter can accomplish a new outcome, which is more attractive to this voter than the original outcome, by means of strategic behavior, by giving the insincere preference $y\ z\ u\ v\ x$, where the Borda winner $x$ is put at the lowest position.

Thus, the Borda procedure gives a voter the possibility to get his or her preferred outcome by giving an insincere order of preference. Hence, on application of the Borda procedure, cheating can be advantageous. The Borda procedure is *not immune to strategic behavior*, or the Borda procedure is *manipulable.*

When Borda was informed of the fact that his procedure was sensitive to strategic behavior, he apparently answered that his procedure was only intended for honest men (Black, 1958, p. 238).

Despite the fact that there are obvious objections to the Borda procedure, it scores relatively well in a comparison of many election procedures (Brams and Fishburn). We would like to quote the following passage from the conclusions of [6]:

'Among ranked positional scoring procedures to elect one candidate, Borda's method is superior in many respects, including susceptibility to strategic manipulation, propensity to elect Condorcet candidates, and ability to minimize paradoxical possibilities. ... Despite Borda's superiority in many respects, it is easier to manipulate than many other procedures. For example, the strategy of ranking the most serious rival of one's favorite candidate last is a transparent way of diminishing the rival's chances.'

'Most votes count' is also sensitive to strategic behavior. This can be seen as follows. For the profile $p$ in Example 1 (page 150), Florence is chosen on application of 'Most votes count'. However, the seven students with individual preference orderings 'Venice Siena Florence' would rather go to Siena than to Florence. This coalition of seven students can accomplish that, on application of 'Most votes count', Siena becomes the collective destination, by giving the insincere individual preference ordering 'Siena Venice Florence'.

The obvious question now is whether the Majority Rule is sensitive to strategic behavior. It can be shown that the possible strategic behavior of a *coalition* $S$, a group of voters, in determining a Condorcet winner would be disadvantageous for at least one of the members of that coalition. So, on application of the Majority Rule (pairwise comparison) for any coalition, there will be at least one member that is disadvantaged due to the strategic behavior of his or her coalition.

**Theorem 3** Suppose $S$ is a coalition. Suppose that profile $p$ renders the true preferences of the voters and that $q$ is the profile in which the individuals in $S$ give insincere preference orderings instead of true preference orderings. Let alternative $x$ be the Condorcet winner for the true profile $p$ and alternative $y$ the Condorcet winner for the insincere profile $q$. Also suppose that $x \neq y$. Then there is an individual $i$ in coalition $S$ that prefers alternative $x$ to $y$. Hence, for that individual, the strategic behavior of the coalition $S$ is disadvantageous, as $y$ will be the Condorcet winner for $q$, while individual $i$ prefers $x$.

**Proof**: Suppose $S$ is a coalition, that is, a (sub)set of individuals. Also suppose that $x$ is the Condorcet winner for the true profile $p$ and that $y$ is the Condorcet winner for the insincere profile $q$, in which only the individuals in $S$ do not give their true preference orderings. Furthermore, suppose that $x \neq y$. Because $x$ is the Condorcet winner for profile $p$, for profile $p$ it holds that $x$ defeats $y$ on pairwise comparison. And because $y$ is the Condorcet winner for $q$, it holds for profile $q$ that $y$ defeats $x$ on pairwise comparison. Hence, there is an individual $i$ such that
1. $i$ prefers $x$ to $y$ for $p$, and
2. $i$ prefers $y$ to $x$ for $q$. (Somebody must have switched preferences.)
As only voters from coalition $S$ give different preference orderings, individual $i$ must be in coalition $S$. Since $i$ prefers $x$ to $y$, $i$ is punished for the strategic behavior of the coalition $S$ to which he or she belongs. □

### 2.7 Approval Voting

Approval Voting assumes that the voter can divide the alternatives into two classes: the candidates that he or she approves of and the ones that he or she disapproves of. The number of candidates that is found to be acceptable can vary, depending on the voter. In the ultimate case, someone can find all alternatives acceptable. The candidate who gets the most votes this way, is the winner. Because the voter mentions all candidates that he approves of, he enlarges the chance that a candidate he finds acceptable will win.

We will divide the acceptable and not acceptable alternatives by means of $\gg$. With, for example,

$$\text{Florence Siena} \gg \text{Venice}$$

we indicate that an individual orders the alternatives from left to right in descending order of acceptability, and that the individual in question only finds Florence and Siena to be acceptable alternatives.

Approval voting is elaborately discussed and propagated by Brams and Fishburn [4].

Now consider the following profile $\hat{p}$, which differs from profile $p$ in Example 1 only in the appearance of the division mark $\gg$.

| | | | | | |
|---|---|---|---|---|---|
| Florence | | Venice | $\gg$ | Siena | : 5 students |
| Florence | | Siena | $\gg$ | Venice | : 7 students |
| Venice | $\gg$ | Florence | | Siena | : 3 students |
| Venice | $\gg$ | Siena | | Florence | : 7 students |
| Siena | | Florence | $\gg$ | Venice | : 3 students |
| Siena | | Venice | $\gg$ | Florence | : 6 students |

Florence then gathers $5 + 7 + 0 + 0 + 3 + 0 = 15$ votes. Siena is good for $0+7+0+0+3+6 = 16$ votes. And Venice now gets $5+0+3+7+0+6 = 21$ votes. So, for this profile $\hat{p}$, Venice is the collective choice under Approval Voting. The collective preference ordering is

$$\text{Venice Siena Florence.}$$

In order to see that the winner under Approval Voting need not necessarily be the Condorcet winner, consider the following profile with three alternatives $a$ (Ann), $b$ (Bob), and $c$ (Coby) and nine voters.

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | : 5 voters |
| $b$ | $a$ | $c$ | : 2 voters |
| $c$ | $b$ | $a$ | : 2 voters |

Then $a$ is the Condorcet winner. Now suppose that, under Approval Voting, all voters give their approval only to the first two alternatives in their respective preference orderings. Then, under Approval Voting, $b$ is the winner, while $a$ is the Condorcet winner.

Approval Voting is also sensitive to strategic behavior. In the example at the beginning of this section, the last group of six students prefers Siena to Venice. Now, by not giving their true preference

$$\text{Siena Venice} \gg \text{Florence}$$

but their insincere preference

$$\text{Siena} \gg \text{Venice Florence},$$

they ensure that Venice gets 6 votes less, $21 - 6 = 15$, and, hence, Siena, with 16 votes, is the collective choice, which is the preferred alternative for these six students.

However, the strategic behavior of one individual or a group of individuals may have the consequence that alternatives which are acceptable to this individual or group get less votes or that unacceptable alternatives get more votes.

We quote from the conclusions of [6]:
'Among non-ranked voting procedures to elect one candidate, approval voting distinguishes itself as more sincere, strategy proof, and likely to elect Condorcet candidates than other procedures ... . Its use in earlier centuries in Europe [...], and its recent adoption by a number of professional societies - including the Institute of Management Sciences [...], the Mathematical Association of America [...], the American Statistical Association [...], the Institute of Electrical and Electronics Engineers [...], and the American Mathematical Society - augurs well for its more widespread use, including possible adoption in public elections [...]. Bills have been introduced in several U.S. state legislatures for its enactment for state primaries, and its consideration has been urged in such countries as Finland [...] and New Zealand [...].'

The reader may easily verify the following theorem.

**Theorem 4** Approval Voting is anonymous, neutral, IIA, transitive, not monotonic and does not satisfy the Pareto condition.

### 2.8 Summary

There are many ways to aggregate individual preferences to a collective preference or outcome. Some of the more frequently occurring ones have been discussed here. For the same individual preferences of the voters, in general, the outcome strongly depends on the election mechanism used.

'*Most votes count*' (Plurality Rule) is very frequently used and is the foundation of the Dutch and British election systems. We have shown that this election mechanism has many disadvantages: it does not satisfy the majority principle (if the number of voters that prefer $x$ to $y$ is greater than the number of voters that prefer $y$ to $x$, then alternative $x$ must also end above $y$ in the outcome), it does not satisfy the Pareto condition (if everybody prefers $x$ to $y$, then $x$ must also be collectively preferred to $y$), and it does not have the monotonicity property (if alternative $x$ is raised vis-a-vis an alternative $y$ in someone's preference ordering and $x$ goes down in no one's preference vis-a-vis $y$, then $x$ must also be raised vis-a-vis $y$ in the outcome). 'Most votes count' can even give an alternative as winner that is defeated by all other alternatives.

The *Majority Rule* (or pairwise comparison) is based on the majority principle. In comparison to 'Most votes count', the Majority Rule has many advantages: not only is it anonymous and neutral, but it is also monotonic and IIA, and it satisfies the Pareto condition. The Majority Rule has just one serious disadvantage: in some situations, it may happen that no winner can be selected, for instance, in the case of three alternatives $x$, $y$, and $z$, where $x$ defeats $y$, $y$ defeats $z$, but also $z$ defeats $x$. In other words, the Majority Rule (pairwise comparison) is not transitive.

The *Borda preference rule* also takes into account the second, third, etc., preferences of individuals in the determination of the collective preference. Frequently, but not always, the Borda preference rule generates the Condorcet winner (if there is one), which is the alternative that defeats all other alternatives on pairwise comparison. The Borda preference rule is not independent of irrelevant alternatives, but perhaps the greatest objection that can be raised against the Borda procedure is its sensitivity to strategic behavior. Nonetheless, the Borda preference rule is, with respect to choosing one single candidate, in many ways superior to other procedures that also weigh second, third, etc., preferences.

*Approval Voting* gives the voter the opportunity to distinguish between the candidates he or she approves of and the ones he or she does not approve of. It is sensitive to strategic behavior of the voter(s). However, among non-ranked voting procedures to elect one candidate, Approval Voting distinguishes itself as more sincere, more strategy proof, and more likely to elect Condorcet candidates than other procedures. However, it is known (see [16]) that the chance of selecting a Condorcet winner, if there is one, under Approval Voting is significantly smaller than under the Borda procedure.

## 3 Categoric voting

Generally, in Western Europe four different election procedures can be distinguished. This is the result of a division in the way of representation (proportional or non-proportional) and in the way of voting (categoric or ordinal). As to the way of representation, we distinguish (see [13], p. 4):

- *Proportional Representation (PR)*: the distribution of seats is proportional to the number of votes.
- *Non-Proportional Representation*: the distribution of seats is not proportional to the number of votes.

Concerning the way of voting we distinguish (see [32], pp. 17, 126):

- *categoric voting*: the voters cast one vote, meaning that they select one candidate or party.
- *ordinal voting*: the voters give a preferential order of candidates or parties. For instance, in Australia, Ireland and Malta the voters are allowed, instead of casting just one vote, to give their first, second, third, etc. preference.

On the basis of the aforementioned distinctions in the way of representation and the way of voting, we can, in general, distinguish four different categories of election procedures, as given in the scheme below.

| | *Way of voting* | |
|---|---|---|
| *Representation* | Categoric | Ordinal |
| Proportional | NL and most European countries | Ireland (STV), Malta (STV) |
| Non- Proportional | UK, US, Canada, New Zealand | Australia (AV), France (two voting rounds) |

Here STV stands for *Single Transferable Vote*, to be considered in section 4.1 and AV for *Alternative Vote*, to be considered in section 4.3. Besides the abovementioned election procedures, there are also the so-called hybrid systems, such as the ‘two-vote’ system in Germany, which we will discuss in section 3.4.

For each category we will discuss a particular election procedure, and we will show the paradoxes this procedure can give rise to. With a paradox, we mean an outcome that is completely contrary to what we would expect or contrary to our sense of righteousness.

In section 3.1 we discuss the main ideas of the Dutch and British election procedures. In section 3.2 a number of paradoxes that may occur in the Dutch system are considered, while in section 3.3 the paradoxes in the British system are discussed.

In sections 4.1 and 4.2 the Single Transferable Vote system is considered. Finally, in sections 4.3 and 4.4 the Alternative Vote system, as applied in, for instance, Australia and to a certain degree also in France, is discussed.

### 3.1 The Netherlands vs The United Kingdom

The Dutch election procedure is characterized, among others, by:

- proportional representation, where the parties receive a number of seats more or less proportional to the number of votes (according to the d'Hondt method; see the end of this section).
- One district, containing the entire nation.

Because of this, there exist more parties and the government usually consists of a coalition of a number of parties.

The British election procedure, on the contrary, is characterized by:

- a division in (approximately) 659 districts for (approximately) 659 seats.
- in each district precisely one representative is elected, by means of 'Most votes count' (Plurality Rule): in each district the party with the most votes wins the seat.

Because of this, the United Kingdom (England, Wales, Scotland and Northern-Ireland) has a two (recently three) party system and the government is usually formed by the party with a majority of the seats.

In general, the Dutch and British election procedures produce different outcomes given the same individual preferences of the voters. To illustrate the difference in outcome between the Dutch and British election procedures, we consider the following example in [29], table 12.1:

| *Party* | *vote percentage* |
|---|---|
| A | 30 |
| B | 25 |
| C | 20 |
| D | 15 |
| E | 10 |

In the Dutch system, every party will get a number of seats more or less proportional to the number of votes. Party A will then get approximately 30% of the seats, part B 25%, etc. Hence, it is to be expected that a multiparty system evolves (five in this example). Because, in general, none of the parties will get a majority of the seats in parliament, government will usually consist of a coalition of several parties.

Now suppose that the same distribution as given in the above-mentioned table occurs in every district in the United Kingdom. Then, in the British system, the seat for each district is given to party A, because this party has most votes in every district. So, in the British system, the other parties would get no seat at all! Because of the nature of the British election procedure, where only large parties have a realistic chance of a seat, and because of the strategic behavior of voters who do not want to waste their vote on a party that has no chance at all, it is to be expected that the British system will give rise to a two- (or three-) party system. This phenomenon is called *Duverger's law*. Since one party usually gets a majority of the seats in parliament, British government usually consists of one party.

Notice that in the Netherlands 'Most votes count' (Plurality Rule) is used to generate a collective (order of) *preference*: A is collectively preferred to B, B to C, etc. Contrary to this, in the United Kingdom 'Most votes count' is used to establish for each district a collective *choice*: the candidate for party A.

As far as appreciation of the Dutch and British election procedures is concerned, Lijphart remarks in [24], page 144:

1. If much weight is given to the representation of minorities, then proportional representation and more than two parties seem to be the best choice.
2. If, on the contrary, much weight is given to government responsibility, then 'Most votes count' (Plurality Rule) and a two-party system seem to be the best choice. The voter then knows that the ruling party is responsible for the achievements of the past and can hold this party responsible for them.

According to [27], pp. 173-175, different considerations concerning representation (democracy) are at the basis of the Dutch and British systems:

1. The Dutch system corresponds to the *reflection model* of representation. The underlying consideration is that the composition of parliament must be a reflection of the composition of the constituency. Various groups and interests are to be proportionally represented, as in a representative test sample. Ideally, proportionally, there will be as many liberals, socialists, etc. in parliament as there are in society.
2. The British system corresponds to the *principal agent model* of representation. According to this model, representatives are agents that act in the interest of others. Parliament does not have to be a reflection of society, but has to honestly defend the interests of the constituency. Not the composition of parliament but its decisions are important.

The *Ostrogorski paradox* shows that the formation of parties and voting for them may give results that deviate from voting for issues, as is done in referenda. Suppose that there are two parties: X and Y. Also, suppose that these two parties have different points of view concerning three issues, numbered 1, 2 and 3. Finally, suppose that there are four groups of voters, named A (20%), B (20%), C (20%) and D (40%), whose positions concerning the three issues are given in the table below, taken from [8], p. 205. For instance, the voters in group A share the position of party X concerning issues 1 and 2, and the position of Y on issue 3.

| *Voters* | *Issues* | | | *Elected party* |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| A (20%) | X | X | Y | X |
| B (20%) | X | Y | X | X |
| C (20%) | Y | X | X | X |
| D (40%) | Y | Y | Y | Y |
| | Y: 60% | Y: 60% | Y: 60 % | |

We now distinguish between two forms of voting:

1. *issue-by-issue voting*: a voting round is held for each separate issue.
2. *voting by platform*: a party or candidate is chosen on the grounds of its policy.

For the situation given in the table, these two forms of voting result in completely different outcomes. If we take issue-by-issue voting, party Y will get 60% of the votes for all issues and will, hence, be able to impose its position on society.

With voting by platform, the voter chooses the party that approximates his or her own position best. The voters in group A will then vote for party X because this party holds their position on two of the three issues. Given this form of voting, party X will get 60% of the votes and party Y only 40%. So now party X has a majority and is able to impose its position concerning the issues on society.

The conclusion is that the outcome of issue-by-issue voting can be completely different from the outcome of voting by platform.

Notice that for three issues that have to be decided on by yes or no, there are $2^3 = 8$ different possibilities to answer these questions. So there would have to be at least 8 different parties to give the voter the possibility to vote for a party that holds his position on all issues. Because in the Netherlands there are more parties than in Britain, the probability of the occurrence of the Ostrogorski paradox will be slightly smaller in the Netherlands.

As has been noticed before, the Netherlands has a system of proportional representation. However, it seldom occurs that the seats can be distributed precisely proportionally to the number of votes. To appoint seats to parties in a more or less proportional manner, the *d'Hondt formula* is used in the Netherlands.

This formula uses the numbers 1, 2, 3, 4, ... to divide the total number of votes a party has received, every time the party gets a seat. The first seat goes to the largest party, whose number of votes is then divided by two. The second seat is allocated to the party that now has the most votes, given that the number of votes the largest party had has now been divided by two. When the largest party receives a second seat, its total number of votes is then divided by three, and so on. The effect of the d'Hondt formula is illustrated by means of the following example from [24], p. 154, in the case of six seats.

| *Party* | $v$ (= votes) | $v/2$ | $v/3$ | *number of seats* |
|---|---|---|---|---|
| A | 41,000 (1) | 20,500 (3) | 13,667 (6) | 3 |
| B | 29,000 (2) | 14,500 (5) | 9,667 | 2 |

| | | | |
|---|---|---|---|
| C | 17,000 (4) | 08,500 | 1 |
| D | 13,000 | | 0 |

### 3.2 Paradoxes in the Dutch system

This section is based on [9] and on [11]. In the first article, Van Deemen shows that a number of paradoxes may occur in the Dutch system. Next, the authors of the second article show, by means of empirical research, that most of these paradoxes do, in fact, occur more than once.

To start with, let us look at the distribution of votes and seats in the elections of September 6, 1989 for the House of Commons, given in the following table. Here SR stands for Small Right, a coalition of some smaller parties.

| *Party* | *Percentage of votes* | *Number of seats* |
|---|---|---|
| CDA | 35.3 | 54 |
| PvdA | 31.9 | 49 |
| VVD | 14.6 | 22 |
| D66 | 7.9 | 12 |
| GL | 4.1 | 6 |
| SR | 5.0 | 7 |

Now consider the following profile, where the distribution of first votes for the parties corresponds precisely to the election results of September 6, 1989. Notice that the profile, though fictitious, is not unrealistic.

| | | | | | | |
|---|---|---|---|---|---|---|
| CDA | D66 | VVD | SR | PvdA | GL | : 35.3% |
| PvdA | GL | D66 | CDA | VVD | SR | : 31.9% |
| VVD | PvdA | D66 | SR | CDA | GL | : 14.6% |
| D66 | PvdA | CDA | VVD | GL | SR | : 07.9% |
| GL | PvdA | D66 | CDA | VVD | SR | : 04.1% |
| SR | VVD | CDA | D66 | PvdA | GL | : 05.0% |

**Paradox 1**: The reader can check for himself that application of pairwise comparison to the above-mentioned profile, yields the following result. Here #(X) stands for the number of seats given to party X.

– PvdA defeats CDA with 58.5% (31.9 + 14.6 + 7.9 + 4.1) to 40.3% (35.3 + 5.0), while #(CDA) = 54 > #(PvdA) = 49.
– VVD defeats PvdA with 54.9 % (35.3 + 14.6 + 5.0) to 43.9% (31.9 + 7.9 + 4.1), while #(PvdA) = 49 > #(VVD) = 22.
– D66 defeats VVD with 79.2% to 19.6%, while #(VVD) = 22 > #(D66) = 12.
– D66 defeats CDA with 59.5% to 40.3%, while #(CDA) = 54 > #(D66) = 12.

Van Deemen calls this the *More-Preferred, Less-Seats paradox*: a party that (in a pairwise comparison) is more preferred than another party may still get fewer seats!

*The Dutch election procedure is also sensitive to strategic behavior.* If for the above profile it is expected, on the grounds of predictions of voting outcomes, that CDA and PvdA will form a coalition, then voters with SR as first preference (the last group of 5%) could choose strategically and mention VVD (their actual second choice) as their (insincere) first choice, hoping to make a coalition of CDA and VVD possible. Such a coalition would then represent 35.3 + 14.6 + 5 = 54.4% of all voters.

Also, if predictions concerning the voting outcome show that one's most preferred party will not make the election threshold (the minimal percentage of votes needed to get a seat, 0.67% in the Netherlands), this voter could decide not to vote for his or her true first preference in order to avoid wasting the vote.

**Paradox 2**: In the elections of September 6, 1989 there was a party, called the Groenen (Greens), that did not get sufficient votes to win a seat. We denote this party with the letter G. Now consider the following profile:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| CDA | G | D66 | VVD | SR | PvdA | GL | : 35.3% |
| PvdA | G | GL | D66 | CDA | VVD | SR | : 31.9% |
| VVD | G | PvdA | D66 | SR | CDA | GL | : 14.6% |
| D66 | G | PvdA | CDA | VVD | GL | SR | : 07.9% |
| GL | G | PvdA | D66 | CDA | VVD | SR | : 04.1% |
| SR | G | VVD | CDA | D66 | PvdA | GL | : 05.0% |

This profile originates from the previous profile by placing party G second in every row. So, it is supposed that every Dutchman has party G as his second preference. The reader can easily check that, given this profile, party G will defeat every other party in a pairwise comparison and, hence, is a Condorcet winner. However, under the Dutch election procedure, party G will get no seat at all!

Van Deemen calls this the *Condorcet-Party-Turns-Loser paradox*: A Condorcet winner does not necessarily get the largest number of seats; it may even happen that the Condorcet winner gets no seat at all.

**Paradox 3**: The following result is even more amazing. Consider the following profile:

| | | | | | | |
|---|---|---|---|---|---|---|
| CDA | GL | SR | D66 | VVD | PvdA | : 35.3% |
| PvdA | GL | SR | D66 | VVD | CDA | : 31.9% |
| VVD | GL | SR | D66 | PvdA | CDA | : 14.6% |
| D66 | GL | SR | VVD | CDA | PvdA | : 07.9% |
| GL | SR | D66 | VVD | PvdA | CDA | : 04.1% |
| SR | GL | D66 | VVD | PvdA | CDA | : 05.0% |

For this profile it holds that in a pairwise comparison

GL defeats SR, SR defeats D66, D66 defeats VVD,
VVD defeats PvdA, and PvdA defeats CDA

in other words, GL has a majority over SR, which has in its turn a majority over D66, etcetera. But this is precisely the inverse of the collective preference as given by the distribution of seats:

$$\#(\text{CDA}) > \#(\text{PvdA}) > \#(\text{VVD}) > \#(\text{D66}) > \#(\text{SR}) > \#(\text{GL})$$

in other words, on the above-mentioned fictitious profile CDA gets more seats than PvdA on application of the Dutch system, PvdA will get more seats than VVD, et cetera.

Van Deemen calls this the *Reversal-of-Majority paradox*: the order given by application of the Majority Rule (pairwise comparison) is exactly the inverse of the order given by the actual distribution of seats in the Dutch system.

Notice that the last of the three paradoxes given is the strongest: an occurrence of the Reversal-of-Majority paradox entails the occurrence of the Condorcet-Party-Turns-Loser paradox; and the latter entails the occurrence of the More-Preferred, Less-Seats paradox.

We cite here A. van Deemen, [9], page 240:
'It is hard to find reasons that justify the possibility that a candidate or party which is preferred by a minority is elected or has more seats than a candidate or party which is preferred by a majority. Borda (1781) and Condorcet (1788) rightly concluded that for this reason the plurality systems are 'seriously defective' ([2], p. 44). The paradoxes presented in this paper lead to the same conclusion for list systems of proportional representation.'

Above, we have constructed, behind our desks, three situations or profiles in which the Dutch system, based on 'Most votes count', leads to paradoxical results. The obvious question now is if such situations also occur in real life. Well then, in [11] the authors describe the results of their empirical research concerning the occurrence of situations (profiles) in Dutch elections that could lead to one of the above-mentioned paradoxes. In short, their findings are that they could not find an occurrence of the strongest paradox, the Reversal-of-Majority paradox, but that the other paradoxes occur frequently.

The reader may wonder how such empirical research is possible, since the voters are only asked to give their first preference. How can we know what their second, third, etcetera preferences are? During the run-up to every election so-called voter research is done, in which a number of voters is asked to give their individual preference ordering with respect to *all* parties. By making the number of participants sufficiently large, reliable information concerning the individual preferences of the voters over all parties can be gathered.

– [11], page 484: For the elections of 1982, 1986 and 1994, the More-Preferred, Less-Seats paradox frequently occurred: *a party that has a majority over an other party can still get fewer seats.* This paradox also occurred in 1989, be it to a lesser extent.

- [11], page 485: The Condorcet-Party-Turns-Loser paradox occurred in the elections of 1982 and 1994: *it can happen that a Condorcet winner does not get the largest number of seats, or perhaps even no seats at all.* In 1994, D66 was Condorcet winner, but PvdA got most of the seats. Also, CDA and VVD got more seats than D66. A second case occurred in 1982 when CDA got more seats than the Condorcet winner PvdA. The possibility of a Condorcet winner getting no seats at all did not occur.
- [11], page 485: the Reversal-of-Majority paradox did not occur in the elections of 1982, 1986, 1989 or 1994.

### 3.3 Paradoxes in the British system

In what follows we will analyze three paradoxes of the British system and pay some attention to the May 1948 election in South-Africa.

*Condorcet-Loser-Wins paradox* In the separate districts, where 'Most votes count' is used, the paradoxes we have seen in the Dutch system will, of course, also occur. Suppose, for instance, that the preferences of the voters in a district with respect to three candidates Ad ($a$), Bob ($b$) and Carol ($c$) are as follows:

$a \quad b \quad c \quad : 30\%$
$b \quad a \quad c \quad : 30\%$
$c \quad b \quad a \quad : 40\%$

For this profile, on application of 'Most votes count' (Plurality Rule), $c$ is elected, while $b$ is the Condorcet winner. Worse, $c$ is the *Condorcet loser*, meaning that all other candidates have a majority over $c$. There is a majority (60%) that prefers $a$ to $c$ and a majority (60%) that prefers $b$ to $c$.

A second paradox in the British system is caused by the division in districts and is, therefore, called the *districts paradox*. Suppose that there are three districts, two parties A and B, twenty voters in each district and that the votes are divided over the candidates for the two parties as follows.

| | *candidate for A* | *candidate for B* | *elected* |
|---|---|---|---|
| *district 1* | 11 votes | 9 votes | A |
| *district 2* | 11 votes | 9 votes | A |
| *district 3* | 5 votes | 15 votes | B |

When 'Most votes count' (Plurality Rule) is applied, the candidate for party A will win in districts 1 and 2, and in district 3 the candidate for party B will win. According to the British system, party A will then have a majority in the House of Commons and, hence, form a government. But B has 33 votes, which is more than the 27 votes for A. So, on direct elections, B would have won and formed the government.

The majority that party A acquires is called, in [32], pp. 74-75, a *manufactured majority*: a majority in the legislative power, won by a party that has got fewer votes than the other party. According to empirical research of [24], page 74, British parliamentary elections over the period 1945-1990 produced manufactured majorities in 92.3 percent of all cases.

A similar situation occurred in the *elections in South-Africa in May 1948*. There were two parties, the United Party, against Apartheid, and the National Party, in favor of Apartheid. The United Party got 50.9%, so more than half of all votes, but due to the district system only 71 seats. The National Party only got 41.2% of all votes, but 79 seats in parliament. This party won the elections due to the single-member district plurality system: a district system where in each district one candidate is chosen by means of 'Most votes count'. This shows again that the choice of the election procedure can have far-reaching consequences.

The *districts paradox* that may occur in the British system is illustrated again by means of the following example from [29], pp. 221-222. Suppose there are five parties A, B, C, D and E, ten districts and in each district 10% of the voters, and the following distribution of votes (in percentages).

| | *DISTRICT* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *PARTY* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| A | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| B | 3 | 3 | 2 | 2 | 3 | 0 | 3 | 3 | 3 | 3 |
| C | 0 | 0 | 0 | 0 | 0 | 4 | 4 | 4 | 4 | 4 |
| D | 0 | 0 | 4 | 4 | 4 | 3 | 0 | 0 | 0 | 0 |
| E | 4 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |

Because party C gets most votes in districts 6 to 10, this party will get 50% of the seats in the British system, even though the party only gets 20% of the votes nationally. Party D wins in the three districts 3, 4 and 5 and so gets 30% of the seats, even though the party only gets 15% of the votes nationally. Party E wins in the districts 1 and 2 and gets 20% of the seats in the British system, while the party only gets 10% of the votes nationally. Party A, on the contrary, that by far gets most votes nationally, 30%, gets no seat at all in the British system.

In the above-mentioned example, parties C, D and E together get all seats in parliament, while only getting 20 + 15 + 10 = 45% of the votes. Parties A and B get no seat in parliament at all, while getting 30 + 25 = 55% of all votes. It is now obvious that parties A and B will dissolve themselves, leaving only three parties. In reality we also see that only 2 to 3 parties are active in Great Britain. Hence, the above paradox illustrates why one can expect that 'Most votes count' (Plurality Rule) combined with a district system will generate a system with few parties (*Duverger's law*). However, as this example shows, it will not necessarily be the small parties that disappear.

In the separate districts 'Most votes count' is used, which - as we already saw - is sensitive to strategic behavior. Therefore also *the British election system is sensitive to strategic behavior*, as is illustrated in the following example. Suppose that in a certain district there are three candidates A (Ann), B (Bob) and C (Cod) and the preferences of the eighteen voters are as follows.

| | | | |
|---|---|---|---|
| A | B | C | : 6 voters |
| C | A | B | : 5 voters |
| B | C | A | : 4 voters |
| B | A | C | : 3 voters |

For 'Most votes count' this means that candidate B will be elected in this district. An opinion poll preceding the elections could cause the five voters with candidate C as most preferred and candidate B as least preferred candidate to change their vote and put candidate A first. The result would then be that A defeats B with 11 to 7 votes and this outcome is preferred to the original outcome by these five voters.

Note that in the British system a party that wins the district's seat after joining two districts, does not necessarily win in both original districts. In order to see this, consider the following two profiles for district 1 and district 2.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| District 1: | $a$ | $b$ | $c$ | : 9 voters | District 2: | $a$ | $b$ | $c$ | : 6 voters |
| | $b$ | $c$ | $a$ | : 5 voters | | $b$ | $c$ | $a$ | : 9 voters |
| | $c$ | $b$ | $a$ | : 3 voters | | $c$ | $b$ | $a$ | : 2 voters |

*The Dutch and British system are not Independent of Irrelevant Alternatives.* In order to see this, consider the following profile.

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | : 4 voters |
| $c$ | $b$ | $a$ | : 3 voters |
| $b$ | $c$ | $a$ | : 2 voters |

1) Check that in the Dutch system the number of seats for $c$ will be greater than the number of seats for $b$, but that on absence of $a$ the inverse will be the case. The collective preference concerning $b$ and $c$, as expressed in the distribution of seats, is, hence, influenced by the presence of the (irrelevant) alternative $a$.
2) Check that in the British system the district's seat is given to neither $b$ nor $c$ and that these two parties are hence collectively indifferent, but that on absence of $a$ the district's seat is given to $b$ and not to $c$. So also in the British system the collective preference, as expressed in the distribution of seats, between $b$ and $c$ is influenced by the presence of the (irrelevant) alternative $a$.

### 3.4 Hybrid systems

A country like *Germany* has a hybrid election procedure, the so-called *two-vote system.* It combines two ideals, namely district representation, as in England, and proportional representation, as in The Netherlands. Every voter has two votes: a first vote (Erststimme) for one candidate from his or her district and a second vote (Zweitstimme) for one of the national parties. Half of the Bundestag consists of district candidates and half of representatives of the national parties. Furthermore, there is an election threshold of 5%. The voter may vote with his first vote for a candidate that does not necessarily belong to the party he or she gives the second vote to. (For reasons of simplicity we do not consider Überhangmandate.)

Similar systems are used in Mexico, South-Korea, Taiwan and Venezuela. Versions of this system have recently been applied in Hungary, Italy, Japan, New Zealand and Russia (see [13], p. 87). In the Netherlands such a two-vote system was proposed in 1995. The proposal was to give each voter two votes. Parliament will then consist of 75 national seats and 75 district seats. The country would be divided into five districts, each of which distributes 15 seats among the district candidates. Every voter's first vote would be for a candidate from his district, the second vote would be for a national party (as in the current system).

A. van Deemen showed that the proposed election procedure could lead to what has been named the *two-vote paradox*: a party B could get a majority over party A in each district, while party A gets a greater number of seats in parliament than party B. To see this, we suppose four parties A, B, C and D and a distribution of the national seats as follows:

| *Party* | *Number of seats* | *Percentage* |
|---|---|---|
| A | 30 | 40,0% |
| B | 20 | 26,7% |
| C | 15 | 20,0% |
| D | 10 | 13,3% |

We also suppose that the district elections lead to the following result.

| *Districts:* | *D1* | *D2* | *D3* | *D4* | *D5* | *Total* |
|---|---|---|---|---|---|---|
| Party A | 5 | 5 | 5 | 4 | 4 | 23 |
| Party B | 6 | 8 | 6 | 6 | 5 | 31 |
| Party C | 2 | 1 | 2 | 2 | 2 | 09 |
| Party D | 2 | 1 | 2 | 3 | 4 | 12 |
| Total: | 15 | 15 | 15 | 15 | 15 | 75 |

For the above-mentioned outcome, party A will get 30 national and 23 district seats, so the party will have 53 seats in parliament. Party B will get 20 national and 31 district seats, so a total of 51. According to the second table, however, party B has a majority over party A in each district.

Because the proposed election procedure is a combination of the British and the current Dutch system, it is also prone to the paradoxes that may occur in either of these systems (see section 3.2 and 3.3).

The proposed system is an amelioration of the existing one in the sense that more information is asked of the voter and processed. The function of an election mechanism is to aggregate information received from the voter to distribute seats. Such an election procedure, however, cannot serve to, for instance, close the gap between voters and politicians.

### 3.5 Summary

The Dutch election mechanism is based on proportional representation and categoric voting. In section 3.2 we have discussed a number of paradoxes that may occur in this system: the More-Preferred, Less-Seats paradox, the Condorcet-Party-Turns-Loser paradox and the Reversal-of-Majority paradox.

The British system is based on non-proportional representation, as well as on categoric voting. In section 3.3 we have found some paradoxes that may occur in this system: the Condorcet-Loser-Wins paradox and the districts paradox.

In the Netherlands 'Most votes count' (Plurality Rule) is used to generate a collective (order of) *preference*. Contrary to this, in the United Kingdom 'Most votes count' is used to establish for each district a collective *choice*: the candidate for a certain party.

The German hybrid election procedure is a combination of the Dutch and British system and, hence, inherits the paradoxes of both systems. The German two-vote system also creates its own paradox, which we have called the two-vote paradox.

## 4 Ordinal Voting

In ordinal voting, voters are asked to list the candidates or parties in order of preference. In section 4.1, the Single Transferable Vote (STV) system, used in Ireland and Malta, is considered. That this system gives rise to a number of very akward paradoxes is made clear in section 4.2. In section 4.3 we consider the Alternative Vote (AV) system that is used in Australia. That this system is also subject to paradoxes is the subject of section 4.4. The French election system and the paradoxes inherent in it are considered in section 4.5.

### 4.1 The Single Transferable Vote (STV) System

In Ireland and Malta, the Single Transferable Vote (STV) system is used. This system was first proposed by Thomas Hare (1861) in England and Carl George Andrae in Denmark around 1850. Hare presented his system as a way to secure the proportional representation of important minorities. His idea was that no vote should be wasted: even if someone's vote does not aid in electing his or her first choice, the vote can still count for his or her lower choices. John Stuart Mill (1862) classified STV 'among the greatest improvements yet made in the theory and practice of government' ([6], 11.1).

The members of parliament are elected per district. In Ireland, there are about 40 districts that have to elect approximately 150 members of parliament. This means that each district must appoint several representatives. Even though STV is a system of proportional representation, it differs from proportional systems with lists (parties) because the voters choose individual candidates instead of parties. Furthermore, they are asked to give an order of preference regarding the candidates, independent of the parties they belong to, instead of casting just one vote for one candidate (or party), so that the voting is ordinal instead of categoric. To gain a seat, the candidate must pass a certain election threshold.

Let us begin by considering an example that will make clear what moved Hare to develop his STV system. The example is taken from [19], page 211. Consider an imaginary district in which two of the four candidates must be elected. Two candidates, Ann ($a$) and Bob ($b$), are conservative, the other two, Coby ($c$) and Donald ($d$), are progressive. Suppose that the preferences of the 23 voters are as follows:

| | | | | |
|---|---|---|---|---|
| $a$ | $b$ | $c$ | $d$ | : 7 voters |
| $b$ | $a$ | $c$ | $d$ | : 6 voters |
| $c$ | $d$ | $b$ | $a$ | : 6 voters |
| $d$ | $c$ | $b$ | $a$ | : 4 voters |

In an election where each voter may vote for two candidates, $a$ and $b$ will win with 13 votes each. Hence, the 10 progressive voters remain unrepresented, even though they constitute 43% of the electorate. The 13 conservative voters get a 100% representation.

Before giving a general description of Hare's STV system, let us use the above example to see how it works.
The *election threshold* in this example is 8, because

1. for 23 voters, there can be no more than 2 candidates that get eight votes of first choice: three candidates with eight votes of first choice would require $3 \times 8 = 24$ voters;
2. with an election threshold of 7, three instead of two candidates could get seven votes of first choice: $3 \times 7 = 21$.

Thus, the *election threshold* is the smallest number of votes of first choice so that the maximum amount of candidates that can reach the election threshold corresponds to the available number of seats.

Because the election threshold is 8, in the above example every candidate is short of votes. The least popular candidate, Donald ($d$), is then (under STV) eliminated and his four supporters then transfer their votes to Coby ($c$), their second choice. In the second round, the list of preferences then looks as follows:

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | : 7 voters |
| $b$ | $a$ | $c$ | : 6 voters |
| $c$ | $b$ | $a$ | : 10 voters |

Now $c$ exceeds the election threshold by two votes and is hence elected. Her remaining votes are transferred to $b$, the second choice of this group. The situation in the third round then becomes:

| | | |
|---|---|---|
| $a$ | $b$ | : 7 |
| $b$ | $a$ | : $6 + 2 = 8$ |

Now $b$ reaches the election threshold and is elected. Note that on application of STV both the progressive candidate $c$ and the conservative candidate $b$ are elected, while without STV the two conservative candidates $a$ and $b$ would have been elected. So, on application of STV, both groups of voters are more or less proportionally represented.

In the above example, in which 2 (of the 4) candidates had to be elected according to the STV procedure by 23 voters, the election threshold was 8.

Now suppose there are $n$ voters and $k$ available seats. The election threshold $q$ is then, by definition, the smallest natural number such that $kq \leq n$ and $(k+1)q > n$. Thus, $q = [\frac{n}{k+1}] + 1$, where $[\frac{n}{k+1}]$ is the integer obtained by rounding down $\frac{n}{k+1}$.

So, in the above example, where 23 voters have to elect 2 candidates, the election threshold $q = [\frac{23}{2+1}] + 1 = 7 + 1 = 8$.

We now give a general *description of STV*, the system of Single Transferable Vote, as found in [19], page 212. If we assume that at least one candidate reaches the election threshold and at least one seat remains, the winning votes that pass the election threshold are proportionally transferred to the second choice of these voters. If as a result of this transfer another candidate reaches the election threshold, this candidate is elected; and if seats remain, the remaining votes (the ones that passed the election threshold) are once again transferred proportionally. This process continues until all seats are occupied. If at any point there is a seat unoccupied without there being votes to be transferred, the candidate with the least number of votes is eliminated and the supporters of this candidate transfer their votes to their most preferred candidate amongst those that are still in the running (that is, not eliminated and not yet occupying a seat).

The idea is that no vote is wasted: each vote beyond the number a candidate needs to be elected must be counted elsewhere; a vote that is wasted on the least popular candidate must be counted elsewhere.

Let us illustrate how STV works by means of a more complex example, taken from [24], page 158. Suppose that one district has to elect three representatives. It has a hundred voters and seven candidates: P, Q, R, S, T, U, and V. Suppose that the individual preference orderings are given in the following profile.

| | |
|---|---|
| P Q R | : 15 voters |
| P R Q | : 15 voters |
| Q R P | : 8 voters |
| R P Q | : 3 voters |
| S T | : 20 voters |
| T S | : 9 voters |
| U | : 17 voters |
| V | : 13 voters |

Hence, the election threshold is $[\frac{100}{3+1}] + 1 = 26$. On application of STV, the following occurs. In the first round, P has 15 + 15 = 30 votes and is the only candidate to reach the election threshold. In the second round the four $(30-26)$ surplus votes P had are proportionally transferred to the second choice of these voters. In this case, two to Q and two to R, because half of the original thirty preferences with P as first choice had Q as second choice and half had R as second choice. In the second round, Q then has 8 + 2 = 10 (first) votes, R has 3 + 2 = 5, S 20, T 9, U 17, and V 13 votes. So, none of these candidates reaches the election threshold in the second round. Hence, in the third round the weakest candidate, R, is eliminated and the five votes R had are transferred to the first candidate that is still in the running, which is Q. In the third round, Q then has 10 + 5 = 15 votes, S 20, T 9, U 17, and V 13 votes. This means that also in the third round, no candidate reaches the election threshold. In the fourth round, according to the STV procedure, again the weakest candidate is eliminated, in this case T, and T's nine votes are transferred to S, the second choice of these voters. In the fourth round, Q has 15 votes, S 20 + 9 = 29, U 17, and V 13 votes. S now passes the election threshold and is elected. At this stage, P and S are elected while R and T have been eliminated. Because three candidates must be elected, a fifth round is necessary: the three (29 − 26) surplus votes S had must be transferred to the next preference of the voters, in this case T. Since T has already been eliminated, the three votes are not transferable. So, nobody reaches the election threshold in the fifth round, and, in the sixth round, the weakest candidate is removed. This is V, whose 13 votes are not transferable. Only Q with 15 and U with 17 votes remain. This means that in the seventh round, the weakest candidate, Q, is eliminated and that U becomes the third candidate elected.

## 4.2 Paradoxes in the STV System

Even though Hare developed his Single Transferable Vote system with the best of intentions, this system frequently leads to results that are contrary to what one would expect. In this section we will expound on a number of these paradoxes.

*STV is majority-inconsistent*: A Condorcet winner may exist without being elected by STV. We have already seen in sections 3.2 and 3.3 that the Dutch and British systems, both based on 'Most votes count', are also majority-inconsistent: the Condorcet-Party-Turns-Loser paradox in the Dutch system and the Condorcet-Loser-Wins paradox in the British system. Probably, the first to recognize the majority-inconsistency of STV were Hoag and Hallett (see [18]).

Suppose that STV is to be used to elect one of three candidates, Ann ($a$), Bob ($b$), and Coby ($c$), by eight voters. The election threshold is then $[\frac{8}{1+1}] + 1 = 5$. Suppose that the individual preferences are given in the following profile:

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | : 3 voter |
| $b$ | $a$ | $c$ | : 2 voters |
| $c$ | $b$ | $a$ | : 3 voters |

Because nobody reaches the election threshold in the first round, STV requires that $b$, being the candidate with the fewest (first) votes, be eliminated and his two votes transferred to the second choice of these voters, $a$. Consequently, $a$ reaches the election threshold with $3 + 2 = 5$ votes and is elected.

However, on application of the Majority Rule (pairwise comparison), we see that $b$ defeats $a$ with $2 + 3 = 5$ to 3 votes and $b$ defeats $c$ with $3 + 2 = 5$ to 3 votes. Thus, $b$ is the Condorcet winner, while $b$ is the first to be eliminated in the STV procedure.

*STV is sensitive to strategic behavior.* One of the instructions for the elections of the American Mathematical Society (AMS) was 'there is no tactical advantage to be gained by marking few candidates.' However, in 1982, Steven Brams showed, by means of an example, that, on application of STV, it may well be beneficial to some voters to mention fewer candidates in their preference ordering. In [6], Example 11.1, he gives the following example.

Suppose that two of the four candidates, $x$, $a$, $b$, and $c$, have to be elected by 17 voters who have the following preferences (divided over three (preference) classes I, II, and III).

| | | | | | |
|---|---|---|---|---|---|
| I | $x$ | $a$ | $b$ | $c$ | : 6 voters |
| II | $x$ | $b$ | $c$ | $a$ | : 6 voters |
| III | $x$ | $c$ | $a$ | $b$ | : 5 voters |

The election threshold for STV is now $[\frac{17}{2+1}] + 1 = 6$. On application of the STV procedure, $x$ and $a$ are elected. This goes as follows. In the first round, $x$, with all 17 votes, is the only one to reach the election threshold and is elected. The $17 - 6 = 11$ surplus votes for $x$ are then transferred in the ratio $6 : 6 : 5$ to the second preferences of the voters in group I, II, and III. Groups I and II each receive $\frac{6}{17} \times 11 = \frac{66}{17}$ votes and III receives $\frac{5}{17} \times 11 = \frac{55}{17}$ votes.

| | | | | |
|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | : $\frac{66}{17}$ votes |
| II | $b$ | $c$ | $a$ | : $\frac{66}{17}$ votes |
| III | $c$ | $a$ | $b$ | : $\frac{55}{17}$ votes |

Since after the transfer of votes no candidate reaches the election threshold, in the third round, the candidate with the fewest votes, $c$, is eliminated, resulting in this preference profile:

| | | | |
|---|---|---|---|
| I | $a$ | $b$ | : $\frac{66}{17}$ |
| II | $b$ | $a$ | : $\frac{66}{17}$ |
| III | $a$ | $b$ | : $\frac{55}{17}$ |

$a$ now has $\frac{66}{17} + \frac{55}{17}$ votes, reaches the election threshold, and is hence elected with $x$.

The voters in class II see their least preferred candidate $a$ elected alongside their most preferred candidate $x$. Now suppose that two of the six class II voters had only given their first choice $x$. The profile would then look like this.

| | | | | | |
|---|---|---|---|---|---|
| I | $x$ | $a$ | $b$ | $c$ | : 6 |
| II.1 | $x$ | | | | : 2 |
| II.2 | $x$ | $b$ | $c$ | $a$ | : 4 |
| III | $x$ | $c$ | $a$ | $b$ | : 5 |

The reader can easily verify that on application of the STV procedure candidate $c$ is now elected alongside $x$ (see also [19], pp. 214-215).

Note that the outcome in which $c$ and $x$ are elected is more attractive to class II voters, as they prefer $c$ to $a$. We have already seen that if all voters in class II had given their complete preference ordering, $a$ would have been elected. So, the two class II voters who gave only their first choice have thereby gained a tactical advantage.

*Negative Responsiveness paradox*: On application of STV, a candidate's position can deteriorate when he gets extra votes; more votes can even turn a winner into a loser! In other words, *STV is not monotonic*. Doron and Kronick have clarified this by means of the following example (see [12]). Suppose that 26 voters must elect two of four candidates: Ann ($a$), Bob ($b$), Coby ($c$), and Donald ($d$). Further, suppose that their individual preferences are given in the following profile.

| | | | | | |
|---|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | $d$ | : 9 voters |
| II | $c$ | $d$ | $b$ | $a$ | : 6 voters |
| III | $d$ | $c$ | $b$ | $a$ | : 2 voters |
| IV | $d$ | $b$ | $c$ | $a$ | : 4 voters |
| V | $b$ | $c$ | $d$ | $a$ | : 5 voters |

As the election threshold is 9, $a$ is elected. In the second round, $b$ is dropped (he has the least votes), after which $c$ is elected with $6 + 5 = 11$ votes.

Now suppose that the two voters in class III come to prefer $c$ over $d$ and leave the rest of their preferences unchanged. Then $c$ gains two votes. The new situation is then:

| | | | | | |
|---|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | $d$ | : 9 |
| II | $c$ | $d$ | $b$ | $a$ | : 6 |
| III* | $c$ | $d$ | $b$ | $a$ | : 2 |
| IV | $d$ | $b$ | $c$ | $a$ | : 4 |
| V | $b$ | $c$ | $d$ | $a$ | : 5 |

Once again, $a$ is elected in the first round. Consequently, $d$ is removed (having the least votes), after which $b$ is elected with 5 + 4 = 9 votes.

In the above example, we see that candidate $c$ turns into a loser by getting two more votes. If the two voters in class III* had ranked $c$ second instead of first, $c$ would have been elected.

*Two-districts paradox* ([19], pp. 220-221): A candidate can win in each separate district and still lose, on application of STV, an election in a combination of those districts.

To see this, consider two districts with 21 voters and the same four candidates Ann ($a$), Bob ($b$), Coby ($c$), and Donald ($d$). Suppose that one candidate must be elected in each district, so that the election threshold in each district is 11. The individual preferences for districts 1 and 2 are given by the following profiles.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | $d$ | : 8 voters | I | $a$ | $b$ | $c$ | $d$ | : 8 voters |
| II | $b$ | $c$ | $d$ | $a$ | : 4 voters | II | $b$ | $c$ | $d$ | $a$ | : 4 voters |
| III | $c$ | $a$ | $d$ | $b$ | : 3 voters | III | $c$ | $a$ | $d$ | $b$ | : 6 voters |
| IV | $d$ | $c$ | $b$ | $a$ | : 6 voters | IV* | $d$ | $a$ | $b$ | $c$ | : 3 voters |

In both districts, STV will elect $a$: originally nobody reaches the threshold, but in the second round $c$ and $d$ will be eliminated, after which $a$ reaches the election threshold in both districts with 8 + 3 = 11 votes.

If the two districts are combined into one district and we assume that the individual preferences of the 42 voters remain the same and that once again one candidate must be elected, the STV procedure has a surprise in store for $a$: STV now elects $c$! The election threshold is now 22. Initially, nobody reaches the election threshold and $b$ is eliminated. Hence, $c$ gets 3 + 6 + 4 + 4 = 17 votes. In the next round, $d$ is eliminated. Six of his votes go to $c$, who gets 17 + 6 = 23 votes and reaches the election threshold.

In their article 'Paradoxes of Preferential Voting; What can go wrong with sophisticated voting systems designed to remedy problems of simpler systems' ([14]), Fishburn and Brams give a beautiful example of what they call the no-show paradox. Because a couple cannot participate in a mayoral election due to the breakdown of their car, they unwittingly prevent their least preferred candidate from winning. Had they been able to vote, their least preferred candidate would have won on application of STV.

*No-show paradox*: By adding individual preferences with $x$ as the least preferred alternative, $x$ can turn into a winner (while originally being a loser).

Fishburn and Brams describe a situation in which a village has to elect a mayor according to the STV procedure, whereby a car breakdown prevents Mr. and Mrs. Smith from participating in the election. Both have the preference ordering: Ann Bob Cod, where they strongly dislike Cod. Owing to their absence, there are only 1608 voters that have to elect one mayor according to the STV procedure. The individual preferences are given in the following profile.

| *Amount* | *Preference* |
|---|---|
| 417 | Ann Bob Cod |
| 82 | Ann Cod Bob |
| 143 | Bob Ann Cod |
| 357 | Bob Cod Ann |
| 285 | Cod Ann Bob |
| 324 | Cod Bob Ann |
| 1608 | |

On application of STV, the election threshold is $[\frac{1608}{1+1}] + 1 = 805$. Ann has 417 + 82 = 499 first votes, Bob 143 + 357 = 500, and Cod has 285 + 324 = 609 first votes. So, in the first round, none of the candidates reaches the election threshold and the candidate with the fewest first votes, Ann, is eliminated. The votes for Ann are then transferred to Bob and Cod. This means that, in the second round, Bob gets 500 + 417 = 917 votes, reaches the election threshold, and is elected mayor. Note that Bob is the second preference of the Smith couple, who are very happy that Cod, whom they despise, did not win.

Now let us see what would have happened had the Smith's car functioned properly and they had been able to participate in the election. The election result would then look like the one above, only with 419 voters with preference 'Ann Bob Cod' and a total of 1610 voters.

Because there are two more voters with the preference 'Ann Bob Cod', Ann has two more first votes: 501 instead of 499. This means that, in the first round, Bob, who received only 500 first votes, is eliminated instead of Ann and that his votes are proportionally transferred to Ann and Cod. Hence, Cod gets 609 + 357 = 966 votes, reaches the election threshold $[\frac{1610}{1+1}] + 1 = 806$, and is elected mayor. Mr. and Mrs. Smith are perturbed: by not showing up (no-show), they unwittingly prevented their least preferred candidate from winning the election.

It is easy to verify that all other paradoxes given in this section can be illustrated by means of the above example:

a) The winner in the above example according to STV is also the Condorcet winner in the case of a car breakdown, but not in the case where the Smith couple was able to participate in the election. (*STV is majority-inconsistent*).

b) Consider the case in which the Smith couple is able to vote. If two or more of the 82 voters with preference 'Ann Cod Bob' put Cod in first place (Cod Ann Bob), then Ann is eliminated first instead of Bob, and Bob wins instead of Cod. An increase in support for Cod turns him into a loser (*Negative Responsiveness paradox*).

c) The Smith couple's village has two districts: East and West. The votes were distributed over the districts as follows.

| Amount | Preference | East | West |
|---|---|---|---|
| 417 | Ann Bob Cod | 160 | 257 |
| 82 | Ann Cod Bob | 0 | 82 |
| 143 | Bob Ann Cod | 143 | 0 |
| 357 | Bob Cod Ann | 0 | 357 |
| 285 | Cod Ann Bob | 0 | 285 |
| 324 | Cod Bob Ann | 285 | 39 |
| 1608 | | 588 | 1020 |

Ann would have won in both separate districts (on application of STV), while she loses in a combined vote for the two districts (*Two-districts paradox*). Note that both Bob and Cod have a majority over Ann on a merging of the districts.

### 4.3 The Alternative Vote (AV) System

In Australia, the members of parliament are chosen per district (as in England), where each district elects one member. As the Australian parliament has 148 seats, the number of districts is 148 (see [24], page 17). The voters are asked to give their preference ordering of the candidates. To aggregate the individual preferences per district to an election of one candidate, the Alternative Vote procedure is used.

With the Alternative Vote (AV) procedure, the candidate who gets more than 50% of the first votes is elected. If no such candidate exists, the candidate with the fewest first votes is eliminated. The votes for that candidate are then transferred to the other candidates in accordance with the second preferences. This procedure is repeated until one of the candidates gets more than half of all votes. In other words, as long as no candidate gets more than 50% of the votes, the candidate with the fewest first votes is eliminated and his votes are transferred in accordance with the second, third, etc. preferences.

The following example illustrates this procedure. Suppose that there are 24 voters and five candidates: Ann ($a$), Bob ($b$), Coby ($c$), Donald ($d$), and Edward ($e$). The individual preferences are given in the following profile.

| | | | | | |
|---|---|---|---|---|---|
| $a$ | $e$ | $c$ | $b$ | $d$ | : 4 voters |
| $b$ | $a$ | $d$ | $c$ | $e$ | : 5 voters |
| $c$ | $d$ | $b$ | $e$ | $a$ | : 8 voters |
| $d$ | $a$ | $e$ | $b$ | $c$ | : 2 voters |
| $d$ | $c$ | $b$ | $e$ | $a$ | : 1 voters |
| $e$ | $a$ | $b$ | $d$ | $c$ | : 2 voters |
| $e$ | $d$ | $b$ | $c$ | $a$ | : 2 voters |

In the first round, nobody has more than half (that is, 12) of the first votes, and candidate $d$ is eliminated with the fewest first votes. Two votes are then transferred to $a$ and one to $c$, the second preferences of the voters with $d$ as first preference. In the second round, $a$ has $4 + 2 = 6$ votes and $c$ has $8 + 1 = 9$. Since still none of the candidates has more than half (12) of all votes (24), candidate $e$ is eliminated with the fewest first votes. Two votes for $e$ are then transferred to $a$, the second preference of these voters, and, because candidate $d$ has already been eliminated, two votes for $e$ are transferred to candidate $b$. In the third round, $a$ then has $6 + 2 = 8$ votes and $b$ has $5 + 2 = 7$ votes.

| | $a$ | $b$ | $c$ | $d$ | $e$ |
|---|---|---|---|---|---|
| number of votes in first round | 4 | 5 | 8 | 3 | 4 |
| second round | 6 | 5 | 9 | - | 4 |
| third round | 8 | 7 | 9 | - | - |
| fourth round | 13 | - | 11 | - | - |

In the third round, still none of the candidates has more that half of the votes and candidate $b$ is eliminated with the fewest (7) votes. Of the seven votes for $b$, five are transferred to $a$, the second choice of the voters with $b$ as first preference. The other two votes for $b$ are transferred to candidate $c$. These are the two votes that candidate $b$ acquired after candidate $e$ was eliminated. So, even the fourth preference of these voters is taken into account. With two remaining candidates $a$ and $c$, $a$ is elected in the fourth round because, with 13 votes, he now has more that half of the votes.

### 4.4 Paradoxes in the AV System

The Alternative Vote system also leads to unexpected outcomes in some situations, as is illustrated by the following paradoxes.

*AV is majority-inconsistent*: A Condorcet winner may exist without being elected by the Alternative Vote (AV) procedure. This was already shown by Hoag and Hallett in [18]. Van Deemen in [9], p. 237, gives the following example concerning one district with nine voters and three candidates Ann ($a$), Bob ($b$), and Coby ($c$).

$a$ $b$ $c$ : 4 voters
$c$ $b$ $a$ : 3 voters
$b$ $c$ $a$ : 2 voters

It is easily seen that, in the above example, $b$ is the Condorcet winner, while the Alternative Vote procedure eliminates $b$ first and then elects $c$.

*AV is sensitive to strategic behavior*: In the above profile, the four voters with preference ordering $a$ $b$ $c$ prefer $b$ over $c$. If these four voters vote strategically by giving the untrue preference ordering $b$ $a$ $c$ instead of their true preference ordering, on application of AV, $b$ will have more than half of the votes and will hence be elected. This outcome is preferred by our four voters over the outcome $c$ that results when they give their true preference.

The above example can also be used to show that *AV is not Independent of Irrelevant Alternatives* (not IIA). To see this, consider the profile that is generated from the above profile by dropping $a$. We then have two profiles that are equal as far as $b$ and $c$ are concerned. But AV applied to the original profile gives $c$ as the outcome, while AV applied to the modified profile gives $b$ as the outcome. The presence of the (irrelevant) alternative $a$ therefore influences the preference between $b$ and $c$ on application of AV. Similarly, this example also shows that *STV is not IIA* .

In the case that only one candidate has to be elected, the election threshold on application of STV equals $[\frac{n}{2}] + 1$, where $n$ is the number of voters. So, when electing one candidate by means of STV, he or she will be elected when he or she gets more than half of the votes. Hence, when only one candidate has to be elected, the Alternative Vote (AV) procedure corresponds to the STV procedure. Therefore, the examples in section 4.2 show that the *no-show paradox* and the *Negative Responsiveness paradox* can also occur on application of the *AV procedure*.

The reader can easily check that the *two-districts paradox*, considered in section 4.2, can also occur in the Alternative Vote (AV) system. To see this, consider again the example used to illustrate the two-districts paradox for the STV system.

The *districts paradox* that can occur in the British system can also occur on application of STV or AV. Each can be illustrated by the same example; see section 3.3. In this example, party A will gain a *manufactured majority* in parliament also on application of STV and AV respectively: a majority of the number of seats gained with a minority of (first) votes.

### 4.5 The French Election System

In Australia, seats are distributed on the basis of absolute majority. This means that the seats in parliament are not distributed proportionally to the number of votes a party managed to get. The structure of voting is ordinal: apart from their first preference, voters can also give their second, third, etc. preferences.

In **France**, a similar election procedure is used, one that combines non-proportional representation with a form of ordinal voting. In France, 555 members of parliament are elected in 555 districts by means of the so-called *majority-plurality rule*. According to [24], p. 18, an (absolute) majority (more than half of all votes) is needed in the first round to gain the seat. When none of the candidates gets an absolute majority of votes in the first round, a second and final election round is organized, where the criterion 'Most votes count' (plurality) is used. In general, only two candidates compete in this second round, because the weakest candidates (those with less than 17% of all votes) are forced to withdraw and other candidates are allowed to withdraw in favor of another candidate from an allied party. The French presidential elections are decided by the *majority-runoff* formula. This formula is comparable to the majority-plurality formula, except that only two candidates may participate in the second round, i.e., the ones that got the most votes in the first round. In addition to France, this formula is also used for direct presidential elections in Portugal and Austria.

*The French election procedure with two rounds is also majority-inconsistent.* We illustrate this paradox by means of an example taken from [19], p. 222. Suppose there are three candidates, $Pro$(gressive), $Cen$(ter), and $Con$(servative), and the preference ordering of the voters is given by the following profile.

| | | | |
|---|---|---|---|
| $Pro$ | $Cen$ | $Con$ | : 49% |
| $Cen$ | ? | ? | : 10% |
| $Con$ | $Cen$ | $Pro$ | : 41% |

$Cen$ defeats $Con$ with 49 + 10 = 59% of the votes and $Cen$ defeats $Pro$ with 10 + 41 = 51% of the votes. So, $Cen$ is the Condorcet winner. However, in the first round, nobody gets more than half of the (first) votes, and $Cen$, with the fewest votes, will have to withdraw. The second round will then be between $Pro$ and $Con$, despite the fact that $Cen$ is the Condorcet winner.

Consider (again) the following example:

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | : 4 voters |
| $c$ | $b$ | $a$ | : 3 voters |
| $b$ | $c$ | $a$ | : 2 voters |

It is easy to see that the French election mechanism is *not IIA and sensitive to strategic behavior.*

In their book *Approval Voting* [4], Brams and Fishburn show that the *Negative Responsiveness paradox* may also occur in the French election procedure. Consider three candidates, Ann ($a$), Bob ($b$), and Coby ($c$), and seventeen voters that have the following preferences.

| | | | | |
|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | : 6 voters |
| II | $c$ | $a$ | $b$ | : 5 voters |
| III | $b$ | $c$ | $a$ | : 4 voters |
| IV | $b$ | $a$ | $c$ | : 2 voters |

Then $a$ and $b$ with six votes each will reach the second round, where $a$ will win from $b$ with 6 + 5 = 11 against 4 + 2 = 6 votes. Now suppose that the voters in class IV promote $a$ from their second to their first preference. This gives rise to the following profile.

| | | | | |
|---|---|---|---|---|
| I | $a$ | $b$ | $c$ | : 6 voters |
| II | $c$ | $a$ | $b$ | : 5 voters |
| III | $b$ | $c$ | $a$ | : 4 voters |
| IV* | $a$ | $b$ | $c$ | : 2 voters |

Then $a$ and $c$ will reach the second round with $6 + 2 = 8$ and 5 votes respectively, where $c$ will win from $a$ with $5 + 4 = 9$ against $6 + 2 = 8$ votes. Greater support for $a$ costs her the victory!

The reader can easily check that the example Fishburn and Brams use to illustrate the no-show paradox on application of STV (see section 4.2) can also be used to illustrate that the *no-show paradox* may occur for presidential elections in France.

### 4.6 Summary

The STV (Single Transferable Vote) system is used in Ireland and Malta. This system is based on proportional representation (per district) and on ordinal voting (meaning that the voters give a preference ordering of candidates or parties). It is a fairly complex procedure that, in spite of all the good intentions (proportional representation), gives rise to a number of paradoxes, as seen in section 4.2: STV is majority-inconsistent, sensitive to strategic behavior, and subject to the Negative Responsiveness paradox. Especially the latter seems damning: more votes for a candidate can cause him to lose his seat. Besides, on application of STV, the no-show paradox can occur: adding individual preferences with $x$ as the least preferred candidate can turn $x$ from a loser into a winner. STV is also prone to the districts paradox and the two-districts paradox.

The Alternative Vote (AV) system that is used in Australia (and implicitly in France) is based on non-proportional representation (per district) and on ordinal voting. This system is likewise not free of paradoxes: all the above mentioned paradoxes occur in this system, as we have seen in section 4.4. Like STV, AV is also not Independent of Irrelevant Alternatives.

In section 4.5, we have shown that the French election procedure is also subject to the paradoxes discussed above.

## 5 Arrow's theorem

In the previous chapters, we have considered several election procedures: 'Most votes count' (Plurality Rule), pairwise comparison (Majority Rule), Borda procedure, Approval Voting, the Dutch and British systems both based on 'Most votes count', the Single Transferable Vote (STV) system, and the Alternative Vote (AV) system. For each of these election procedures, we have ascertained a number of properties that we consider to be unwanted or to have negative consequences. Of course, the question arises as to whether any 'good' election procedures exist. However, what are 'good' election procedures?

We have already seen a number of requirements which we could say an election procedure needs to satisfy: anonymity, neutrality, IIA, Pareto condition, monotonicity, not sensitive to strategic behaviour, non dictatorial (which means that there is no individual - called the *dictator* - such that the election procedure selects the preference of that individual in all cases). Furthermore, an election procedure should be *transitive*, meaning that, if the procedure prefers $x$ to $y$ and $y$ to $z$, then the procedure will also prefer $x$ to $z$. We have seen previously that the Majority Rule does not have this property: for the Condorcet profile in section 2.3, we have seen that, on pairwise comparison, Florence defeats Venice, Venice defeats Siena, but Florence does not defeat Siena (on the contrary, Siena defeats Florence). 'Most votes count' is a transitive election procedure: if $x$ has more votes than $y$ and $y$ has more votes than $z$, then $x$ has more votes than $z$.

The question remains if all these properties are unquestionably positive. For instance, anonimity (of the voters) can be considered a great good, but it is sometimes irritating that the vote of a specialist is counted only as heavily (or lightly) as that of the ignorant. On the other hand, if an election procedure is manipulable (meaning that it is sensitive to strategic behaviour), that can be considered a negative property. However, if the possibility of manipulation is small and the procedure has many other positive properties, we might still consider it a 'good' election procedure.

In order to choose from the multitude of available election procedures, it seems wise to determine the properties of the procedures. Based on these properties, positive and negative, it will then hopefully be possible - keeping in mind the purpose of the procedure to be selected - to make a well-considered choice. Even better, but also more difficult, would be to characterise several procedures by means of a number of properties. This means that you show that a procedure with such and such properties must of necessity be one particular procedure. In [25], K. May gave the following characterisation of pairwise comparison (Majority Rule): Majority Rule (pairwise comparison) is the only election procedure that is anonymous, neutral, and monotonic.

There are deep mathematical-logical results at the base of the perpetual wonderment when studying election procedures for more than three candidates ([6], Introduction).

The first observation, by Kenneth Arrow around 1950, is that, for three or more alternatives, there can not be an election procedure that satisfies a number of conditions which could be considered desirable properties for such a procedure. Arrow's theorem states that, for at least three alternatives, there can not be a transitive election procedure - more precisely, there is no transitive (collective) preference rule - that at the same time satisfies the Pareto condition, is Independent of Irrelevant Alternatives (IIA), and non-dictatorial. In other words, K. J. Arrow proved that every transitive election procedure (more precisely, every transitive preference rule) that satisfies the Pareto condition and is Independent of Irrelevant Alternatives (IIA) must necessarily be dictatorial. With this, he gives a characterisation of dictatorial election procedures (preference rules). As a dictatorial election procedure is generally considered to be unwanted, Arrow's result is also called an *impossibility theorem*: a transitive election procedure (preference rule) can never at the same time satisfy the Pareto condition, be Independent of Irrelevant Alternatives, and be non-dictatorial. In other words, the Pareto condition, IIA, and non-dictatoriality are incompatible; it is impossible to construct a transitive election procedure (preference rule) that has all three of these properties. This is quite surprising, as each of these three properties appears to be reasonable at first sight. Arrow's theorem shows us that reaching a collective decision is much more complicated than it would seem.

The second observation, by Gibbard (1973) and Satterthwaite (1975), says that all reasonable election procedures - more precisely, choice rules - for three or more alternatives are sensitive to strategic behaviour. The Gibbard-Satterthwaite theorem states that, for at least three alternatives, there can not be an election procedure (choice rule) that is at the same time Pareto optimal, not sensitive to strategic behaviour, and non-dictatorial. In other words, for at least three alternatives, every Pareto optimal and non-manipulable choice rule is dictatorial.

Notice that Arrow's impossibility theorem is about preference rules, whereas that of Gibbard and Satterthwaite is about choice rules. As we have seen in the Introduction, a *choice rule* selects one alternative for every voter profile, while a *preference rule* assigns to each voter profile a collective order of preference of the alternatives.

A choice rule $K$ is called Pareto optimal if it selects for every profile $p$ an alternative $y$ $(= K(p))$ that is not collectively improvable (meaning that there does not exist an alternative $x$ that everyone prefers to $y$).

A choice rule $K$ is *non-manipulable* or *not sensitive to strategic behaviour* if cheating does not pay, or more precisely, if strategic behaviour of one of the individuals is not beneficiary to that individual independent of what the other individuals do.

The reader interested in the proofs of these theorems, is referred to [10].

Notice that, in both Arrow's theorem and that of Gibbard and Satterthwaite, the condition is 'at least three alternatives'. For two alternatives, 'Most votes count' and the Borda rule coincide with the Majority Rule, and hence are, like the Majority Rule, neutral, anonymous, monotonic, IIA, non-dictatorial and satisfy the Pareto condition.

The challenge posed by the two above mentioned impossibility theorems is not to devise a perfect election procedure. The theorems put forward by Arrow and Gibbard and Satterthwaite show that this is impossible. The challenge is to identify those procedures that aggregate the wishes of the voters to a collective choice or preference (outcome) as loyally as possible to the voter preferences. We would want an election procedure that ([6], Introduction)

- encourages sincere voting (based on true preferences)
- is relatively immune to strategic manipulation
- avoids obvious paradoxes, like the Negative responsiveness paradox, that occurs when increased support for a candidate turns it into a loser where it previously was a winner.

Furthermore, both impossibility theorems assume that all possible individual preference orderings (all possible profiles) are allowed, while, in practice, the voter profiles will often be subject to several restrictions. For instance, someone who has a person from the right-wing as favorite politician is unlikely to have someone from the extreme left-wing as second choice. Under some restrictions, for instance, single peakedness of the profiles, some nice (that is, satisfying a number of desired properties) election procedures (more precisely, choice rules) are possible; see [1] or [34], Chapter 4. Single peakedness of a profile roughly means that the individual preferences in the profile can be ordered along a line in such a way that, if we follow the line from left to right, the preference of each individual grows to a peak and then diminishes. For instance, the individual preferences regarding the temperature in a room are of the following form: every individual has one optimal temperature and, as one deviates from this temperature, the preference will diminish.

The next example is taken from [34], pp. 97-99. Suppose that, at an office, five persons work in the same space, for which the central heating can be set between 15 and 30 degrees Celcius (no decimals). The employees have to decide collectively on the temperature. It is very likely that every individual will have one optimal temperature and that the preference will diminish as one deviates from this optimum. Now consider the profile below, in which the preference concerning the temperature for each of the employees is given. Call this profile $p$. It is an example of a single peaked profile.

preference of individual

measure of preference 1 2 3,4 5

15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

alternative temperatures

The optimum of individual 1 is 15 degrees, that of individual 2 is 20 degrees, 3 and 4 have 23 degrees as optimum, and individual 5 has 30 degrees as optimum. Notice that three of the five individuals, namely 3, 4 and 5, (strictly) prefer 23 degrees to all temperatures $x$ lower than 23 degrees. In other words, for all $x \in \{15, 16, \ldots, 22\}$, 23 defeats $x$ on pairwise comparison, given profile $p$. Notice also that four of the five individuals, namely 1, 2, 3 and 4, (strictly) prefer 23 degrees to all temperatures $x$ higher than 23 degrees. In other words, for each $x \in \{24, 25, \ldots, 30\}$, 23 defeats $x$ on pairwise comparison, given profile $p$.

Hence, given the above (single peaked) profile $p$, 23 defeats every other alternative on pairwise comparison. In other words, 23 is the Condorcet winner given profile $p$.

In general, it can be shown that, for single peaked profiles, there is always a Condorcet winner if there is an uneven number of individuals. The choice rule that assigns to those profiles the Condorcet winner is not sensitive to strategic behaviour (non-manipulable), Pareto optimal, and anonymous. This is also the only choice rule for these profiles that has these properties. See [1], [34].

### 5.1 The Impossibility theorems

In this section, we will give a mathematically precise formulation of the impossibility theorems of Arrow and of Gibbard and Satterthwaite. For proofs of these theorems, we refer to [10] or to [34], Chapter 3. We start with Arrow's theorem.

**Definition 17** Given a set $N$ of individuals, a set $A$ of alternatives, and a (collective) preference rule $F : L(A)^N \to C(A)$, we call $i$ a *dictator* for $F$ if, for every profile $p \in L(A)^N$, $F(p) = p(i)$.
A (collective) preference rule $F$ is called *dictatorial* if there is an individual $i$ such that $i$ is a dictator for $F$.

Hence, a collective preference rule $F$ is non dictatorial if, for every individual $i$, there exists a profile $p$ such that the collective preference $F(p)$ is not equal to the individual preference $p(i)$ of $i$ ($F(p) \neq p(i)$).

It seems reasonable to suppose that the collective preference must be transitive: for all alternatives $x$, $y$, and $z$, if the community prefers $x$ to $y$ and $y$ to $z$, then the community will also prefer $x$ to $z$. A collective preference rule $F$ that generates, for every profile $p$, a transitive collective preference $F(p)$, is called transitive.

**Definition 18** A preference rule $F : L(A)^N \to C(A)$ is called *transitive* if, for every profile $p$ in $L(A)^N$, $F(p)$ is a transitive relation. So, $F : L(A)^N \to C(A)$ is transitive if $F : L(A)^N \to W(A)$, where $W(A)$ is the set of all weak orderings (meaning, complete and transitive) on $A$.

**Theorem 5** *Arrow's impossibility theorem*: Suppose that there are at least three alternatives in $A$ and $F : L(A)^N \to W(A)$ is a transitive (collective) preference rule. If $F$ satisfies the Pareto condition and is Independent of Irrelevant Alternatives (IIA), then $F$ is dictatorial.

In conclusion, we present a mathematically precise formulation of the Gibbard and Satterthwaite theorem.

**Definition 19** A choice rule $K : L(A)^N \to A$ is called *Pareto optimal* if, for every profile $p \in L(A)^N$, there exists no $x$ in $A$ with $x \neq K(p)$ such that for all $i \in N$, $x \; p(i) \; K(p)$. In other words, a choice rule $K$ is called Pareto optimal if it assigns to each profile $p$ an alternative $K(p)$ that can not be collectively ameliorated.

**Definition 20** A choice rule $K : L(A)^N \to A$ is called *dictatorial* if there exists an individual $i \in N$ such that, for each profile $p \in L(A)^N$, the collective choice $K(p)$ is the best alternative in $p(i)$.

A choice rule $K : L(A)^N \to A$ is called *non-manipulable* or *not sensitive to strategic behaviour* if cheating does not pay, or more precisely, if strategic behaviour of an individual is not beneficiary to that individual, independent of what other individuals do.

**Definition 21** $K : L(A)^N \to A$ is *non-manipulable* if, for all $i \in N$ and for all $p, q \in L(A)^N$, if $p(j) = q(j)$ for all $j \neq i$, then $K(p)\ p(i)\ K(q)$. Therefore, $K$ is non-manipulable if, for every profile $p$, unilateral deviation of an individual $i$ from $p$ to $q$ is not beneficiary to $i$ ($K(p)$ is at least as good as $K(q)$ for $i$).

**Theorem 6** *Gibbard/Satterthwaite's impossibility theorem*: Suppose there are at least three alternatives in $A$. Let $K : L(A)^N \to A$ be a Pareto optimal and non-manipulable choice rule. Then $K$ is dictatorial.

### 5.2 Other Characterization Theorems

The theorem of Arrow gives a characterization of the dictatorial preference rule: it is the only transitive preference rule that is IIA and satisfies the Pareto condition (if there are at least three alternatives).

Similarly, the Gibbard/Satterthwaite theorem gives a characterization of the dictatorial choice rule: it is the only choice rule that is Pareto optimal and non-manipulable (if there are at least three alternatives).

In order to decide which are the better election procedures it may be useful to have characterizations of the different procedures. We mention a few below.

In [35] H.P. Young gave a characterization of the Borda choice correspondence: it is the only choice correspondence that is neutral, consistent, faithful and has the cancellation property. The *consistency* condition relates choices made by disjoint subsets of voters to choices made by their union. It says that if two disjoint subsets of voters choose the same alternatives, using a choice correspondence, then their union should choose exactly the same alternatives, using this same choice correspondence. *Faithfulness* demands of a choice correspondence that, if society consists of a single individual, it must choose the most preferred alternative of this individual. A third desirable property introduced by Young is the cancellation property. A choice correspondence has the *cancellation property* if it declares a tie between all alternatives iff for all pairs $(p,\ q)$ of alternatives, the number of voters who prefer $p$ to $q$ equals the number of voters preferring $q$ to $p$.

Recently, Eliora van der Hout ([20]) gave a characterization of the Plurality ranking rule, on which the Dutch elections are based: it is the only preference rule that is consistent, faithful and has the FS-cancellation property. In the context of a social preference rule $F$, *consistency* demands that if two disjoint sets of voters $I$ and $J$ both socially prefer party $p$ to party $q$, using $F$, then their union should also socially prefer party $p$ to party $q$, using $F$. Similarly, it requires that if party $p$ is socially preferred to party $q$ by voter set $I$, using $F$, and voter set $J$ is socially indifferent between party $p$ and party $q$, using $F$, then party $p$ should also be socially preferred by the union of $I$ and $J$, using $F$. A social preference rule is *faithful* if, in case society consists of a single individual who's most preferred party is party $p$, it orders this party $p$ first. A social preference rule is said to have the first score (FS) *cancellation* property if it declares a tie between party $p$ and party $q$ in case the number of individuals who prefer party $p$ most (order $p$ first) equals the number of individuals who prefer party $q$ most (order $q$ first).

In [21] Eliora van der Hout gave a characterization of the British (so called, First Past The Post) election system.

Finally, Rob Bosch gave in [3] a characterization of $k$-vote rules. These rules are like Approval Voting, but now everyone has to approve of a fixed number $k$ of candidates.

## References

1. D. Black, *The Theory of Committees and Elections.* Cambridge University Press, Cambridge, 1958.
2. J.-C. de Borda, Mémoire sur les Elections au Scrutin, 1781. English translation by A. De Grazia, Mathematical Derivation of an Election System. *Isis*, 44: 42-51, 1953.
3. R. Bosch, Characterization of $k$-vote rules. KMA and Tilburg University, Faculty of Philosophy. Submitted for Publication.
4. S. J. Brams and P. C. Fishburn, *Approval Voting.* Birkhäuser, Boston, 1983.
5. S. J. Brams, *Rational Politics: Decisions, Games, and Strategy.* CQ Press, 1985; reprinted by Academic Press, 1989.
6. Steven J. Brams and Peter C. Fishburn, Voting Procedures. In: Kenneth Arrow, Amartya Sen and Kotaro Suzumura (eds.), *Handbook of Social Choice and Welfare.* Elsevier Science, Amsterdam, 2002.
7. Condorcet, *Sur les Elections et autres Textes.* Corpus des Oeuvres de Philosophie en Langue Francaise. Librarie Artheme Fayard, Paris, 1986.
8. H. Daudt, De politieke toekomst van de verzorgingsstaat. In: J.J.A. van Doorn and C.J.M. Schuyt (red.), *De stagnerende verzorgingsstaat.* Meppel, 1978.
9. A. van Deemen, Paradoxes of Voting in List Systems of Proportional Representation. *Electoral Studies*, 12: 234-241, 1993.
10. A. van Deemen, *Coalition Formation and Social Choice.* Kluwer, Dordrecht, 1997.
11. A. van Deemen and N. Vergunst, Empirical evidence of paradoxes of voting in Dutch elections. *Public Choice*, 97: 475-490, 1998.
12. G. Doron and R. Kronick, Single transferable vote: An example of a perverse social choice function. *American Journal of Political Science*, 21: 303-311, 1977.
13. D.M. Farrell, *Comparing Electoral Systems.* Prentice Hall/Harvester Wheatsheaf, 1997.
14. P. C. Fishburn and S. J. Brams, Paradoxes of preferential voting. *Mathematics Magazine* 56: 207-214, 1983.
15. W. V. Gehrlein, The expected probability of Condorcet's paradox. *Economics Letters* 7: 33-37, 1981.
16. W.V. Gehrlein, Condorcet's paradox and the Condorcet efficiency of voting rules. *Mathematica Japonica*, 45: 173-199, 1997.
17. Melvin J. Hinich and Michael C. Munger, *Analytical Politics.* Cambridge University Press, 1997.

18. C. G. Hoag and G. H. Hallett, *Proportional Representation*, Macmillan, New York, 1926.
19. P. Hoffman, *Archimedes' Revenge: the joys and perils of mathematics*. Norton, New York, 1988.
20. E. van der Hout, H. de Swart and A. ter Veer, Axioms Characterizing the Plurality Ranking Rule. Tilburg University, Faculty of Philosophy. Submitted for publication.
21. E. van der Hout, Axioms for FPTP systems. Tilburg University, Faculty of Philosophy. Submitted for Publication.
22. Paul E. Johnson, *Social Choice; Theory and Research.* SAGE Publications, London, 1998.
23. A. Lijphart, *Democracies.* Yale University Press, New Heaven and London, 1984.
24. A. Lijphart, *Electoral Systems and Party Systems: A study of 27 Democracies* 1945-1990. Oxford University Press, Oxford, 1994.
25. K. O. May, A set of independent, necessary and sufficient conditions for simple majority decision. *Econometrica* 20, 680-684, 1952.
26. I. McLean, The Borda and Condorcet Principles: Three Medieval applications. *Social Choice and Welfare* 7, 99-108, 1990.
27. I. McLean, Forms of Representation and Systems of Voting. In: D. Held (ed.), *Political Theory Today*, Polity Press, Cambridge, 1991.
28. I. McLean and A. Urken (Eds.), *Classics of Social Choice*. The University of Michigan Press, 1995.
29. D. C. Mueller, *Public Choice II: A revised edition of Public Choice*. Cambridge University Press, Cambridge, 1989.
30. Hannu Nurmi, *Rational Behaviour and the Design of Institutions: Concepts, Theories and Models*. Edward Elgar, 1998.
31. Hannu Nurmi, *Voting Paradoxes and How to Deal with Them*. Springer-Verlag, 1999.
32. D. W. Rae, *The Political Consequences of Electoral Laws*, Revised Edition. Yale University Press, New Haven, Conn., 1971.
33. Kenneth A. Shepsle and Mark S. Bonchek, *Analyzing Politics: Rationality, Behavior, and Institutions*. W.W. Norton, 1997.
34. A.J.A. Storcken and H.C.M. de Swart, *Verkiezingen, Agenda's en Manipulatie*. Epsilon Uitgaven, Utrecht, 1992.
35. Young, H. P., An Axiomatization of Borda's Rule, *Journal of Economic Theory* 9: 43-52, 1974.