LUNDS UNIVERSITET
Nationalekonomiska institutionen

MASTER'S THESIS, 30 ECTS

On the Strategy-proof Social Choice of Fixed-sized Subsets

AUTHOR: ALEXANDER REFFGEN
SUPERVISOR: LARS-GUNNAR SVENSSON
SEPTEMBER, 2006

Contents

Abstract

This thesis gives a contribution to strategy-proof social choice theory, in which one investigates to what extent there exist voting procedures that never can be manipulated in the sense that some voter by misrepresentation of his preferences can change the outcome of the voting and obtain an alternative he prefers to that honest voting would give. When exactly one element should be elected from a set of at least three alternatives, then the fundamental result in strategy-proof social choice theory, the Gibbard-Satterthwaite theorem, shows that there in general exists no satisfactory non-manipulable voting procedure. However, in many practical voting situations, e.g., when the available alternatives can be ordered on a political left–right scale, individual preferences have a structure which is known as single-peakedness, and in this case it is possible to find reasonable strategy-proof voting procedures.

In this thesis, we analyze the more general voting situation when the number of alternatives that should be elected is greater than one but fixed, which for instance is the case in elections to national parliaments, and we are able to prove results analogous to the single-valued case: in general, there exist no reasonable non-manipulable voting procedures, but when preferences are single-peaked, voting can be made strategy-proof. In connection with our analysis of the strategy-proof social choice of fixed-sized subsets, we obtain also two additional interesting results: firstly, we show that the Gibbard-Satterthwaite theorem not only holds for complete preferences, but also for a large class of domains of partial preferences, and secondly, we are able to make the statement of the original Gibbard-Satterthwaite theorem more precise by proving that every reasonable voting procedure not only can be manipulated, but some voter can manipulate it in such a way that he obtains at least his second best alternative.

Keywords: Strategy-proofness, Multi-valuedness, Gibbard-Satterthwaite theorem, Linked domains, Partial preference relations.

Acknowledgement

I am extremely grateful to my supervisor professor Lars-Gunnar Svensson, both in general, for introducing me to the fascinating field of social choice theory, and in particular, for leading my thoughts into the right ways during my work with thesis. I would also like to thank all who have enriched my life during my work with this thesis, and in particular, my gratitude is directed to my parents Wilfried and Christa Reffgen, Siegfried Reffgen, Moa Leonhardt, Eskil Rydhe, Vidar Holje, and Kristina Rydenfelt. I would like to dedicate this thesis to my friend Malin Fredriksson.

1 Introduction

EVERYONE WHO HAS ever participated in a voting among several alternatives should recognize the following line of thought: It is true that x is my most preferred alternative, but I understand that it has no support among other voters; instead of wasting my vote, I should therefore vote for y, which is an acceptable alternative with a true chance to win, because I do not want to see z being elected. If you do not vote according to your true opinion, we will say that you *misrepresent* your preference, and if your misrepresentation indeed changes the outcome of the voting in a for you beneficial way, we say that you *manipulate* the voting procedure.[1] Of course, whether a voter is able to manipulate a certain voting depends on the way in which the voting is carried out, and since manipulation seems ethically unappealing, one may ask whether there exist any *strategy-proof* voting procedures, i.e., voting procedures that can never be manipulated. Unfortunately, it turns out that if a single element should be elected from a set of at least three alternatives, then there exists no reasonable strategy-proof voting procedure. This result is known as the *Gibbard-Satterthwaite theorem*, and since it has been established in 1973, much research has been devoted to investigate whether its pessimistic conclusion still holds in other voting situations with different assumptions, and in some cases it was indeed possible to find strategy-proof voting procedures (examples of this can be found in Example 2.8, Chapter 3, and Section 5.4 in this thesis). To our knowledge, however, there exists still a voting situation that has not yet been considered in the context of strategy-proofness, namely:

> ***The social choice of fixed-sized subsets:*** *A group of voters has to choose a fixed number of elements from a set of alternatives.*

This voting situation is common in any democracy, the foremost example of course being elections to national parliaments, but it is also present whenever society elects a fixed number of members to a committee. The purpose of this thesis is to *investigate whether there exist any reasonable strategy-proof voting procedures for the social choice of fixed-sized subsets*. In the following, we will propose the specific questions we thereby have to answer.

By definition, a voting procedure is strategy-proof if no voter by misrepresentation can obtain an outcome that he prefers to the outcome that honest voting would give. Thus, to decide whether there exist any strategy-proof voting procedures for

[1] This introduction is kept rather informal in order to convey a better first understanding for the subject of this thesis. Definitions and results will get their precise formulations from Chapter 2 on.

the social choice of fixed-sized subsets, we need first to know when a voter prefers one outcome to another. In the context of the original Gibbard-Satterthwaite theorem, this is straight forward because there, different outcomes of a voting can be ranked directly by voters' preferences; by this, we simply mean that if you by misrepresentation can obtain x instead of y, then this will make you better off if and only if you prefer x to y. This is trivial, of course, but when outcomes are subsets of a fixed size, it is no longer obvious how voters rank different outcomes. Suppose, for example, that a committee has to choose two of the alternatives in the set $\{a_1, a_2, a_3, a_4\}$, and assume that you prefer a_1 to a_2, a_2 to a_3, and a_3 to a_4. If you by misrepresentation can obtain $\{a_1, a_3\}$ instead of $\{a_2, a_3\}$, you should certainly be better off, because besides the common alternative a_3, the former set contains a_1, which you prefer to a_2 in the latter set. But suppose now that misrepresentation can give you $\{a_1, a_4\}$ instead of $\{a_2, a_3\}$. The first set contains your most preferred, but also your worst alternative, whereas the second set contains the two middle alternatives, and it is not clear which set you will prefer. In some cases we will thus be able to draw conclusions about voters' preferences over subsets, but in other cases we will fail to do so. The first question to answer in this thesis must therefore be:

> ***Question 1:*** *When can a voter be assumed to prefer one subset to another, or, more formally, what structure do voters' preferences over subsets of a fixed size have?*

Once we know the structure of voters' preferences over subsets, we are able to define what is meant by manipulation of a voting procedure for the social choice of fixed-sized subsets, and we can turn to the next question:

> ***Question 2:*** *Do there exist any reasonable voting procedures for the social choice of fixed-sized subsets that can never be manipulated?*

We will show that the answer to this question is negative in general, precisely as in the case of the Gibbard-Satterthwaite theorem. However, the general impossibility result of the Gibbard-Satterthwaite theorem can be modified in many real voting situations because voters' preferences have often a structure that makes strategy-proof voting possible. This is, for example, the case when parties participating in a political election can be ordered on a traditional left-right scale. Then preferences have a structure which is known as *single-peakedness*, and it turns out that when society has to elect a single alternative and voters have single-peaked preferences, then there exist reasonable strategy-proof voting procedures. Since single-peaked preferences appear frequently in applications, our analysis of whether strategy-proof social choice of fixed-sized subsets is possible would not be complete if we would not investigate whether we can draw a similar conclusion for the social choice of fixed-sized subsets; thus, we will have to answer the following question:

> ***Question 3:*** *Do there exist any strategy-proof voting procedures for the social choice of fixed-sized subsets if voters' preferences over the available alternatives are single-peaked?*

When the three questions above are answered, the purpose of this thesis is fulfilled. However, in research one should always be grateful when one's analysis of a problem not only solves this problem, but also allows to answer another question of interest, which at first glance may seem unrelated to the original problem. In this thesis, the analysis of the social choice of fixed-sized subsets leads with no extra effort also to a more informative variant of the Gibbard-Satterthwaite theorem. In its original form, the Gibbard-Satterthwaite theorem is a purely qualitative theorem, because it only states that every voting procedure can be manipulated at some instance. But one can imagine that different voting procedures are manipulable to different extents. For instance, one voting procedure maybe allows some voter to obtain his seventh best alternative instead of his eighth best by misrepresentation, whereas another voting procedure is more vulnerable to misrepresentation and some voter can by insincere voting obtain his second best alternative instead of his third best. However, for every voting procedure there must obviously be a highest rank that can be obtained by manipulation, and therefore one may ask the following question, the answer of which emerged in connection with our analysis of Question 2:

> ***Question 4:*** *What is the best alternative that can be obtained at every voting procedure by means of manipulation?*

This thesis is organized as follows: Chapter 2 provides the general background. We start with a brief survey of *social choice theory*, the branch of economics to which this thesis belongs, but then we consider more closely the issue of *strategy-proofness* and discuss the Gibbard-Satterthwaite theorem.

Chapter 3 and Chapter 4 contain the general notions and results needed to answer the questions proposed in this introduction. More concretely, in Chapter 3, we consider *restricted preference domains*, i.e., domains of preferences that arise when preferences, for some reasons, can be assumed to have a certain structure, and it will turn out that some of these domains admit strategy-proof voting procedure, whereas others do not. We will need restricted preference domains in two ways: firstly, preferences over subsets have, as indicated above, a certain structure, and they will thus constitute a restricted preference domain; secondly, also the single-peaked preferences in Question 3 constitute a restricted preference domain.

In Chapter 4, we introduce appropriate notions to describe the structure of preferences over subsets, and we will also show that preference domains that satisfy a certain general condition do not admit strategy-proof voting procedures, a result which we will need in order to answer Question 2.

In Chapter 5, we are then sufficiently prepared to analyze whether strategy-proof social choice of fixed-sized subsets is possible, and we will answer the questions proposed above. Chapter 5 contains also a survey of related literature.

Chapter 6, finally, summarizes the results of this thesis.

This thesis also contains two appendices: Appendix A explains the mathematical notations and techniques used in the formalizations and proofs in this thesis, and it is therefore a good starting point for the mathematically inexperienced reader. Appendix B contains a complete and elementary proof of the Gibbard-Satterthwaite theorem, the main result of strategy-proof social choice theory.

Throughout this thesis, we use a larger number of notations and terms with a very precise meaning, and therefore we found it appropriate to provide both a list of notations and an index at the end of the thesis.

2 Social Choice Theory and Strategy-proofness

THIS THESIS GIVES a contribution to strategy-proof social choice theory. In order to introduce the reader to this branch of economics and to clarify the theoretical background of this thesis, we present in this chapter an overview over the main issues and results in social choice theory in general and in strategy-proof social choice theory in particular.

2.1 What is Social Choice Theory?

The founder of social choice theory and Nobel Laureate Kenneth Arrow begins his classical monograph *Social Choice and Individual Values* with the fundamental observation that "in a capitalist democracy there are essentially two methods by which social choices can be made: voting, typically used to make 'political' decisions, and the market mechanism, typically used to make 'economic' decisions" (1963, 1). Political decision making consists of course also of debate, negotiation, and compromise, for example, but the final decision is in deed most often made by voting. While the market mechanism is studied in traditional microeconomics, the decision making by voting is systematically studied in the field of social choice theory. A voting can be carried out in different ways, which is illustrated by the following two examples, and it is therefore of interest to study voting procedures theoretically.

Example 2.1. The probably simplest voting procedure is the *(ordinary) majority rule*, where every voter has exactly one vote, which he can cast on one of the available alternatives, and the alternative that get most votes will be elected. To be well-defined, this method must be supplemented by an appropriate tie-breaking rule, which for example can be drawing of lots. □

Example 2.2. A more sophisticated voting procedure than the majority rule is the *Borda count*. This procedure takes the voters' entire ranking of the available alternatives into account by allowing the voters to assign points to every alternative. For instance, when there are three alternatives, then voters are allowed to assign three points to one alternative, two points to a second, and one point to the remaining alternative, and the alternative that get most points in total will be elected. □

The majority rule and the Borda count are only two thinkable voting procedures. In fact, the number of all possible voting procedures is enormous. For example, when three voters have to choose one of three alternatives by voting, then there are approximately 10^{103} different voting procedures that can be used to carry out this voting,[2] which should be compared with the number of particles in the universe which sometimes is claimed to be 10^{80}. Different voting procedures can of course lead to different social outcomes, which has consequences for the members in the society, and the choice of voting procedure should therefore be carried out with carefulness. In social choice theory, one specifies hence criterions that a voting procedure preferably should satisfy, and analyzes then which voting procedures actually satisfy these criterions.

In the remainder of this section, we present the main result in social choice theory, *Arrow's theorem*, which shows that the possibility to construct voting procedures with desirable properties has strong theoretical limitations. But first, we introduce the basic assumptions and notations used in this thesis to formalize the analysis of voting procedures: We will consider a society consisting of N individuals, and we will use the set $\mathcal{I} = \{1, 2, \ldots, N\}$ to index these individuals. The society is facing a set of alternatives, denoted by $\mathscr{A}$, which contains M elements and from which one alternative must be chosen.[3] We will assume that the individuals in the society have *preferences* over the alternatives in $\mathscr{A}$, which will be denoted by the letter P, often equipped with an index or primes, and disregarded from Chapter 4, we will assume that preferences satisfy the following properties:

Completeness. A preference P can rank any pair of alternatives, which means that if a and b are two distinct alternatives in $\mathscr{A}$, then P either prefers a to b, or b to a, and in case a is preferred to b by P, we shortly write aPb.

Antisymmetry. A preference P is strict in the sense that if P prefers a to b, then it cannot prefer b to a. Note that this assumption excludes the possibility of indifference between two alternatives.[4]

Transitivity. If a preference P prefers a to b and b to c, then it must also prefer a to c. This assumption makes P's ordering of the alternatives in $\mathscr{A}$ internally consistent.

[2] This number can be derived in the following way: Each voter can choose his top alternative in three possible ways, after which his second alternative must be one of the two remaining ones, and the last alternative is then fixed. Hence, each voter can rank the three alternatives in $3 \cdot 2 \cdot 1 = 6$ ways. Since there are three voters, a voting procedure must thus be able to generate an outcome for $6 \cdot 6 \cdot 6 = 216$ different preference profiles. Choosing one of the three possible alternatives for each preference profile gives then $3^{216} \approx 10^{103}$ different voting procedures.

[3] Both the number of individuals N and the number of alternatives M are assumed to be finite.

[4] If there are infinitely many alternatives available, then this assumption is too restrictive, but it seems justifiable in the present context.

A preference P satisfying these three properties ranks the alternatives in $\mathscr{A}$ in a strict order, and we will use the notation $r_k(P)$ to refer to the alternative that is ranked on the kth place by P. In particular, $r_1(P)$ is that alternative in $\mathscr{A}$ which is preferred to all other alternatives by P, and it is also called the *top alternative* of P. The set of all preferences over the alternatives in $\mathscr{A}$ satisfying the properties above will be referred to as the *set of unrestricted preferences* over $\mathscr{A}$ and will be denoted by the Greek letter Σ. The preference of individual i will be denoted by P_i, and we denote further by $\mathscr{P} = (P_1, P_2, \ldots, P_N)$ the collection of all individual preferences, and call $\mathscr{P}$ a *preference profile*. We will frequently investigate how a social choice is affected when individual i changes his preferences, and thereby, it will be convenient to denote preference profiles by

$$(P_i, P_{-i}) = (P_1, P_2, \ldots, P_{i-1}, P_i, P_{i+1}, \ldots, P_N),$$

where P_{-i} thus stands for the preference profile of all individuals in the society apart from individual i. So far, society was somewhat imprecisely supposed to make its choice using some *voting procedure*. From now on, we will be more specific and assume that social choice is made using a *social choice function* $f : \Sigma^N \to \mathscr{A}$ that assigns to every preference profile a unique alternative, the *social choice*. If $a \in \mathscr{A}$ is the social choice for the preference profile $\mathscr{P} = (P_1, P_2, \ldots, P_N)$, we write $f(P_1, P_2, \ldots, P_N) = a$, or $f(P_i, P_{-i}) = a$, or just $f(\mathscr{P}) = a$.

The first rigorous analysis of to what extent it is possible to aggregate individual preferences in a desirable way was carried out by Kenneth Arrow in the monograph *Social Choice and Individual Values*, which was published in 1951 and which can be seen as the starting-point of social choice theory. For theoretical reasons, Arrow did not study social choice functions, but *social welfare functions*, that is, functions of the form $F : \Sigma^N \to \Sigma$ that aggregate preference profiles to a social preference over the alternatives in $\mathscr{A}$.[5] By definition, a social welfare function is required to assign to *every* preference profile a *unique* social preference, and in addition, Arrow requires a social welfare function to have the following three properties:[6]

Pareto optimality. A social welfare function should respect unanimity in the society in the sense that if all individuals prefer a to b, then also the socially chosen

[5] Arrow's social welfare functions are actually more general than presented here since he allows individual preferences to be *weak*, i.e., preferences do not need to satisfy antisymmetry. However, to keep the exposition simple, we assume throughout this thesis that preferences are antisymmetric.

[6] Arrow's conditions for a social welfare function exist in different variants in the literature, and the exposition here follows closely Mas-Colell et al. (1995, 792–96). For an extensive discussion of Arrow's conditions, we refer to Arrow (1963, 22–33).

preference prefers a to b. Formally, this means that if $(P_1, P_2, \dots, P_N) \in \Sigma^N$ is such that $a\,P_i\,b$ for all $i \in \mathcal{I}$ and $F(P_1, P_2, \dots, P_N) = \bar{P}$, then also $a\,\bar{P}\,b$.

Independence of irrelevant alternatives. A social welfare function should rank any pair of alternatives $a, b \in \mathscr{A}$ only depending on how the individuals in the society rank a and b, but *not* on how they rank other alternatives. Formally, this means that if $\mathscr{P}$ and $\mathscr{P}'$ are two preference profiles that rank a and b equally in the sense that $a\,P_i\,b$ if and only if $a\,P_i'\,b$ for all $i \in \mathcal{I}$, then $a\,F(\mathscr{P})\,b$ if and only if $a\,F(\mathscr{P}')\,b$.

No dictatorship. A social welfare function should not be *dictatorial*, that is, there should be no individual that alone decides on the socially chosen preference. Formally, this means there should be no $i \in \mathcal{I}$ such that $F(P_1, P_2, \dots, P_N) = P_i$ for all preference profiles $(P_1, P_2, \dots, P_N) \in \Sigma^N$.

These conditions may seem to be weak and uncontroversial requirements, but they are nevertheless incompatible, which Arrow showed in his impossibility theorem:

Theorem 2.1 (Arrow's Impossibility Theorem). *Suppose that $\mathscr{A}$ is a finite set of at least three alternatives. Then every social welfare function $F : \Sigma^N \to \Sigma$ that satisfies the conditions of Pareto optimality and independence of irrelevant alternatives is dictatorial.*

Note that Arrow's analysis consists both of a *normative* part, in which he formulates his conditions for a social welfare function, and a *positive* part, where he shows the incompatibility of the required conditions. We will not present a proof of Theorem 2.1,[7] but we illustrate by an example that even a widely accepted voting procedure as the majority rule may lead to unsatisfactory outcomes.

Example 2.3 (The Condorcet Paradox). Consider a small society consisting of three individuals that want to order the alternatives in the set $\mathscr{A} = \{a, b, c\}$ using the majority rule, i.e., one alternative will be socially preferred to another alternative if and only if at least two individuals prefer the former to the latter. Note that the majority rule is well-defined for all preference profiles, it is non-dictatorial, and it satisfies the conditions of Pareto optimality and independence of irrelevant alternatives. Suppose now that the three individuals have the preferences P_1, P_2 respectively P_3, which order the alternatives in $\mathscr{A}$ according to

$$a\,P_1\,b\,P_1\,c, \quad b\,P_2\,c\,P_2\,a, \quad \text{respectively} \quad c\,P_3\,a\,P_3\,b.$$

[7] For a formal proof of Arrow's impossibility theorem, we refer to Arrow (1963, 51–59) or Mas-Colell et al. (1995, 796–799).

Applying the majority rule, we find that the society prefers a to b, b to c, but also c to a. Hence, the majority rule leads in this case not to a transitive social preference, and is therefore neither able to rank the alternatives in $\mathscr{A}$ in an unambiguous order. $\square$

In many practical voting situations, society is of course not primarily interested in obtaining a social ranking of all available alternatives, but wants only to choose one of the alternatives using a social choice function. This seems to be a simpler problem, but it turns out that an analysis similar to that for social welfare functions also can be carried out for social choice functions. For example, Mas-Colell et al. (1995, 807–8) formulate the following three conditions for social choice functions:

Monotonicity. A social choice function $f : \Sigma^N \to \mathscr{A}$ should react in the right way when the members in the society change their preferences in the sense that if an individual moves an alternative up in his preference order, then this should not worsen that alternatives chance to be elected. Formally, this means that if $\mathscr{P}$ and $\mathscr{P}'$ are two preference profiles such that $f(\mathscr{P}) = a$ and $a P_i b$ implies $a P_i' b$ for all $i \in \mathcal{I}$ and $b \in \mathscr{A}$, then we should also have $f(\mathscr{P}') = a$.

Pareto optimality. A social choice function $f : \Sigma^N \to \mathscr{A}$ should respect unanimity in the society in the sense that if all members prefer a to b, then the social choice should not be b, i.e., if $a P_i b$ for all $i \in \mathcal{I}$, then we should have $f(\mathscr{P}) \neq b$.

No dictatorship. A social choice function should not have a *dictator*, that is, an individual that alone decides on the social choice. Formally, this means that there should be no $i \in \mathcal{I}$ such that $f(P_1, P_2, \ldots, P_N) = r_1(P_i)$ for all preference profiles $(P_1, P_2, \ldots, P_N) \in \Sigma^N$.

In analogy with Arrow's theorem, we have the following impossibility result for social choice functions:[8]

Theorem 2.2. *Suppose that $\mathscr{A}$ is a finite set of at least three alternatives. Then every social choice function $f : \Sigma^N \to \mathscr{A}$ that satisfies the conditions of monotonicity and Pareto optimality is dictatorial.*

Arrow's theorem and Theorem 2.2 are strong theoretical limitations for the possibility to carry out a voting in a satisfactory way, but it must be pointed out, as Mas-Colell et al. (1995, 799) does, that "it would be too facile to conclude from it that 'democracy is impossible'. What it shows is somewhat else—that we should

[8] In fact, Theorem 2.2 is not an independent result, but Mas-Colell et al. (1995, 809–11) show that it is an almost immediate consequence of Arrow's theorem.

not expect a collectivity of individuals to behave with the kind of coherence that we may hope from an individual."

This section provided only a short introduction to the main issues of social choice theory. For a readable and more general introduction to the problems of social choice theory, we refer to the Nobel lecture "The possibility of social choice" by Amartya Sen (1999), which also contains an extensive reference list over the immense literature in this branch. A formal and rigorous survey of the fundamental concepts and results in social choice theory can be found in the fifth part of Mas-Colell et al. (1995), and for an overview over present research interest in social choice theory, see Bossert and Weymark (2006).

2.2 Strategy-proof Social Choice Theory

The demands of monotonicity and Pareto optimality we made on social choice functions in the previous section are of course somewhat arbitrary, and one can therefore ask whether there exist other desirable properties that can be satisfied by social choice functions. In the sub-branch of social choice theory that is known as *strategy-proof social choice theory,* one investigates under what conditions social choice functions can be *strategy-proof* in the sense that voters never can gain from misrepresenting their preferences and hence have no incentives to vote tactically.

Below, we explain to what extent strategy-proof social choice is possible, but first, we consider some normative motivations for why strategy-proofness can be regarded a desirable property of a social choice function. In general, tactical voting appears ethically unappealing to many people because in a voting, society assigns one vote to every voter, and one can find that a voter who misrepresents his preferences in order to gain from this tries to take more influence than he has a right to.

Moreover, strategy-proofness simplifies voting for the rational voter since any kind of misrepresentation can give at most the same outcome as sincere voting. In particular, in strategy-proof votings, voters need not worry about whether honest voting may be disadvantageous to them.

Strategy-proofness can thus be a normatively desirable property on its own, but it can also be a means to achieve other goals: Often, the construction of a specific voting procedure is based on some normative desires. If, for example, the constituting assembly in a country is of the opinion that the national parliament should be a miniature copy of the whole population, then a proportional voting system, by which a party's share of seats in the parliament approximately equals the party's share of votes in the elections, is a good choice, and it is implemented for instance in Sweden. On the other hand, if it, in order to guarantee political stability, is desirable that the winning party in the elections is able to govern on its own, then some form of plurality voting, where the country is divided into electoral districts and the candidate that receives most votes in a district takes a seat in the parliament, should be adopted, and such a voting procedure is for example used in the United Kingdom. The realization of such normative goals may, however, depend on whether the individuals in the society vote sincerely, and the designer of a voting procedure can only be sure that it works in the intended way if it is strategy-proof.

Example 2.4. To see that the lack of strategy-proofness can have far-reaching consequences, consider the following situation: Suppose that the elections to the national parliament in a country are carried out using a proportional voting system with a threshold of four per cent, which means that parties that receive less than four per cent of the votes take no seats in the parliament at all. Assume now that a small party is the top alternative of five per cent of the voters. This information is of course not available to anyone, so when it comes to elections, supporters of this party may fell unsure about whether it will clear the threshold, and at least some voters will therefore vote for another party since they do not want to waste their votes. In the end, this may lead to that the party gets less than four per cents of the votes. Hence, the lack of strategy-proofness can cause people to vote tactically, which, as in this case, even can be disadvantageous to them, and we note that the voting procedure above does not work in the intended way because a party with a support of five per cent should belong to the parliament. □

On the other hand, the importance of strategy-proofness should not be overemphasized, because one can object as Sen (1970, 195) that in elections "individuals are guided not so much by maximization of expected utility, but by something much simpler, viz., just a desire to record one's true preference."

We will now formalize the analysis of strategy-proof social choice, starting with the following definition:

Definition 2.1 (Manipulability and Strategy-proofness). A social choice function $f : \Sigma^N \to \mathscr{A}$ is said to be *manipulable* if there for some $i \in \mathcal{I}$ exist $P_i, P_i' \in \Sigma$ and $P_{-i} \in \Sigma^{N-1}$ such that

$$f(P_i', P_{-i}) \ P_i \ f(P_i, P_{-i}) \tag{2.1}$$

If f is not manipulable, we say that f is *strategy-proof.*

Note that manipulation not only means that an individual misrepresents his preference, but that he also changes the social choice and gains from doing so. We illustrate by an example.

Example 2.5. Suppose that two individuals want to choose one of the alternatives in the set $\mathscr{A} = \{a,b,c,d\}$ using the Borda count, i.e., they assign four points to their first, three points to their second, two points to their third, and one point to their fourth alternative, and the alternative with most points in total will be elected. Assume that the two individuals have the preferences P_1 respectively P_2, defined by

$$a\,P_1\,b\,P_1\,c\,P_1\,d \qquad \text{and} \qquad b\,P_2\,c\,P_2\,a\,P_2\,d.$$

If both individuals vote sincerely, alternative b will be elected with seven points. However, if individual 1 pretends that his preference instead of P_1 is P_1', defined by $a\,P_1'\,d\,P_1'\,c\,P_1'\,b$, then alternative b gets only five points and alternative a will be elected with six points. Thus, by misrepresentation, individual 1 can change the social choice and obtain his top alternative instead of his second alternative. □

On the other hand, it is also possible to find voting procedures that are strategy-proof, which the following three examples show:

Example 2.6. Consider a society with an odd number of individuals and suppose that the set of alternatives consists of the two alternatives a and b. In this case, the majority rule is well-defined and obviously strategy-proof because if you prefer a to b, then the only way to misrepresent your preference is to cast your vote on b, but this will never make you better off. □

Example 2.7. Suppose now that the set of alternatives is $\mathscr{A} = \{a,b,c\}$, but that the social choice function $f : \Sigma^N \to \mathscr{A}$ is such that $f(\mathscr{P}) = a$ for *all* preference profiles $\mathscr{P} \in \Sigma^N$. Also this social choice function is obviously strategy-proof because no individual can ever change the social choice by misrepresentation. □

Example 2.8. Consider finally the following voting procedure which involves an element of chance and works for any number of alternatives: The individuals in the society are asked to write down their top alternative on ballots, and the social choice is determined by drawing one of the ballots at random. Also this voting procedure is strategy-proof because if your ballot is drawn, you are best off if you have reported your true top alternative, but if your ballot is not drawn, it does not matter what alternative you have reported.[9] □

The voting procedures in these three examples have of course some shortcomings: The majority rule for two alternatives in Example 2.6 is certainly strategy-proof, but it is not obvious whether it is strategy-proof when the number of alternatives

[9]This example has been taken from Gibbard (1973, 593).

becomes larger than two. Next, the constant social choice function in Example 2.7 is not efficient because even if all individuals have the same top alternative, the social choice can nevertheless differ from this alternative. Finally, the voting procedure from Example 2.8 can seem unsatisfactory since it leaves too much to chance, and even alternatives with little support can be elected. However, if we go beyond these limitations and consider voting procedures that (1) work for at least three alternatives, (2) satisfy unanimity in the society in the sense that if all individuals agree on the same top alternative, then this alternative will also be the social choice,[10] and (3) whose outcome only depends on the preferences of the individuals in the society, and not, for example, on chance, which means that the voting is carried out using a social choice function, then we obtain again an impossibility result:

Theorem 2.3 (The Gibbard-Satterthwaite Theorem). *Let $\mathscr{A}$ be a finite set of at least three alternatives, and suppose that $f : \Sigma^N \to \mathscr{A}$ is a social choice function that satisfies unanimity. Then f is strategy-proof if and only if f is dictatorial.*

The Gibbard-Satterthwaite theorem, proved independently by Gibbard (1973) and Satterthwaite (1975), is the fundamental result in strategy-proof social choice theory, and because of its importance, we present a formal proof in Appendix B. In the following, we will briefly discuss the significance of the Gibbard-Satterthwaite theorem. Note first that the interesting part of the theorem is not that a dictatorial social choice function is strategy-proof, which is quite obvious, but it is the other implication, namely that every non-dictatorial social choice function that satisfies unanimity is manipulable, which is important. A direct consequence of this is, for example, that the ordinary majority rule and the Borda count are manipulable, provided that the number of alternatives is at least three. Note next that the Gibbard-Satterthwaite theorem only implies that a non-dictatorial social choice function is manipulable at *some* preference profile, but it does not tell us at which preference profiles this is the case, nor in which way voters have to misrepresent their preferences in order to be better off. In practice, most social choice functions are actually non-manipulable at most preference profiles, and in secrete votes, it is in general impossible for voters to know whether they are in a position to manipulate a voting. However, when a social choice function is manipulable, it is often enough for a voter to have a vague understanding of other voters' preferences in order to know how to use his vote in the most beneficial way. Note also that Theorem 2.3 only says that some individual at some preference profile can be better off if he misrepresents his preference, but it contains no information about how much an individual can gain

[10] Formally, we say that a social choice function $f : \Sigma^N \to \mathscr{A}$ satisfies *unanimity* if $f(P_1, P_2, \ldots, P_N) = a$ whenever $r_1(P_i) = a$ for all $i \in \mathcal{I}$.

from misrepresentation. In Chapter 4, however, we show that every non-dictatorial social choice function that satisfies unanimity can be manipulated in such a way that some individual obtains at least his second best alternative.

Finally, we would like to remark that strategy-proof social choice theory is not unrelated to the issues of social choice theory considered in the previous section, as it might appear at first glance. In fact, the following two lemmas show that a strategy-proof social choice function that satisfies unanimity also satisfies monotonicity and Pareto optimality. The formulation and proofs of Lemma 2.4 and 2.5 follow closely the exposition in Svensson (1999).

Lemma 2.4 (Monotonicity). *If $f : \Sigma^N \to \mathscr{A}$ is a strategy-proof social choice function, then f satisfies monotonicity.*

Proof. We must show that if $\mathscr{P}$ and $\mathscr{P}'$ are preference profiles such that $f(\mathscr{P}) = a$ and $a\,P_i\,b$ implies $a\,P_i'\,b$ for all $i \in \mathcal{I}$ and $b \in \mathscr{A}$, then also $f(\mathscr{P}') = a$. To this end, suppose first that only individual 1 changes his preference. We argue by contradiction and assume that $f(\mathscr{P}') = a'$ and $a' \neq a$. If $a'\,P_1\,a$, then individual 1 can manipulate f by going from P_1 to P_1', and hence, we must have $a\,P_1\,a'$. But then we must also have $a\,P_1'\,a'$, and individual 1 can manipulate f by going from P_1' to P_1. Thus, the assumption $a' \neq a$ must have been wrong, and we conclude that $f(\mathscr{P}') = a$. The lemma follows now when we change the preferences only for individual 2, individual 3, and so forth. $\square$

Lemma 2.5 (Pareto Optimality). *If $f : \Sigma^N \to \mathscr{A}$ is a strategy-proof social choice function that satisfies unanimity, then f satisfies Pareto optimality.*

Proof. We have to show that if a and b are two distinct alternatives in $\mathscr{A}$ and $(P_1, P_2, \dots, P_N) \in \Sigma^N$ is a preference profile such that $a\,P_i\,b$ for all $i \in \mathcal{I}$, then we have $f(P_1, P_2, \dots, P_N) \neq b$. We argue by contradiction and assume that $f(P_1, P_2, \dots, P_N) = b$. Replacing $(P_1, P_2, \dots, P_N)$ by a preference profile $(P_1', P_2', \dots, P_N')$ with the property that $r_1(P_i') = a$ and $r_2(P_i') = b$ for all $i \in \mathcal{I}$, we get by monotonicity that also $f(P_1', P_2', \dots, P_N') = b$. On the other hand, since $r_1(P_i') = a$ for all $i \in \mathcal{I}$, unanimity implies $f(P_1', P_2', \dots, P_N') = a$, which is a contradiction. Hence $f(P_1, P_2, \dots, P_N) \neq b$, and the lemma is proved. $\square$

In the light of these two lemmas, the Gibbard-Satterthwaite theorem can be seen as a direct consequence of Theorem 2.2, but Lemma 2.4 and 2.5 play also an important role in the direct proof of the Gibbard-Satterthwaite theorem presented in Appendix B. Lemma 2.4 and 2.5 are presented here with their proofs because in Chapter 4 we will need to derive weaker variants of these two lemmas.

3 Restricted Preference Domains

IN THIS CHAPTER, we begin to introduce the theoretical tools we will need in order to investigate to what extent strategy-proof social choice of fixed-sized subsets is possible. In many voting situations, it is reasonable to assume that voters' preferences have a certain structure, and some structures turn out to admit non-dictatorial strategy-proof social choice functions, whereas others do not. In this thesis, such structures, or *restricted preference domains*, come up in two ways: Firstly, in Chapter 5, we will analyze the structure of preferences over fixed-sized subsets, and we will need a criterion to decide whether this structure admits non-dictatorial strategy-proof social choice functions. Secondly, in many practical votings, e.g., in political elections, it is reasonable to assume that voters' preferences have a structure that is known as *single-peakedness*. These preferences are of special interest because they admit non-dictatorial strategy-proof social choice functions when society must choose one alternative, and therefore, we proposed Question 3 in the introduction and asked whether the same is true for the social choice of fixed-sized subsets. This chapter is disposed as follows: Section 3.1 provides a brief introduction to the concept of restricted preference domains. In Section 3.2, we introduce single-peaked preferences and demonstrate the existence of non-dictatorial strategy-proof social choice functions on this kind of preference domains. In Section 3.3, finally, we consider the notion of *linked domains*, which in a modified form will be applied to the structure of preferences over subset of a fixed size.

3.1 Restricted Preferences and Strategy-proof Social Choice

One of the assumptions in the Gibbard-Satterthwaite theorem is that the social choice function has the form $f : \Sigma^N \to \mathscr{A}$, which means that f is defined for *all* possible preference profiles, and the individuals in the society are thus implicitly assumed to be capable of having any of the preferences in Σ as their true preference. Often, however, this may be an unnecessarily general assumption, because one can argue as Sen (1970, 165) that "individual preferences are determined not by turning a roulette wheel over all possible alternatives, but by certain specific social, economic, political and cultural forces", and "this may easily produce patterns in the set of individual preferences". As a consequence of this, we may in a given context have reason to classify some preferences as unreasonable and therefore exclude them from the set of admissible preferences. This is of interest because it will allow us to construct non-dictatorial strategy-proof social choice functions. To see this, recall that a social choice function f by definition is manipulable at the preference profile (P_i, P_{-i}) if

$$f(P_i', P_{-i}) \;\; P_i \;\; f(P_i, P_{-i}) \tag{3.1}$$

for some preference P_i'. If we now, for some objective reasons, can assume that no individual in the society has P_i as his true preference, then f can only theoretically be manipulated at (P_i, P_{-i}), but in practice, this will never happen. If we can exclude sufficiently many of the preferences in Σ, so that there no longer exists a preferences profile (P_i, P_{-i}) satisfying (3.1), then we end up with a social choice function that obviously is strategy-proof.[11] We illustrate this by a simple example.

Example 3.1. Suppose a society has to choose a macroeconomic policy that affects the rate of inflation and the rate of unemployment (and only these). There are three policies available, and they are known to lead to the following outcomes:

	Inflation	Unemployment
Policy A	2 %	4 %
Policy B	4 %	2 %
Policy C	4 %	4 %

Since the number of alternatives is greater than two, there are according to the Gibbard-Satterthwaite theorem no non-dictatorial strategy-proof social choice functions that the society can use to make its decision. But take a closer look on the three policies. It seems reasonable that the individuals in the society at a given rate of unemployment will prefer a lower rate of inflation to a higher rate, i.e., they will prefer Policy A to Policy C. Similarly, at a given rate of inflation, a lower rate of unemployment seems more desirable than a higher rate, so the individuals in the society can be assumed to prefer Policy B to Policy C. On the other hand, the members in the society may disagree on whether low inflation or low unemployment is more important, so an individual may prefer Policy A to Policy B, or vice versa. Consider now the set Σ of unrestricted preferences over the three policies, which contains the following six preferences:

P_1	P_2	P_3	P_4	P_5	P_6
A	B	A	B	C	C
B	A	C	C	A	B
C	C	B	A	B	A

[11] Note, however, that one is not allowed to exclude a preference without reason, but one needs an acceptable economic motivation for doing so.

According to the discussion above, only P_1 and P_2 will appear as true preferences among the members in the society. But now we can easily find a strategy-proof social choice function. For instance, the majority rule, which chooses Policy A if most voters report P_1 and Policy B if most voters report P_2, is strategy-proof, non-dictatorial, and satisfies unanimity. □

The set of all preferences that are admissible in a certain context will in general be a strict subset Ω of Σ, and Ω will then be said to be a *restricted (preference) domain*. In Example 3.1, for instance, we have $\Omega = \{P_1, P_2\}$. Precisely as in the case of unrestricted preferences, we will say that a social choice function $f : \Omega^N \to \mathscr{A}$ is *manipulable* if there exist $P_i, P_i' \in \Omega$ and $P_{-i} \in \Omega^{N-1}$ such that (3.1) holds, but note that all preferences in (3.1) are now assumed to belong to Ω. The restricted preference domain in Example 3.1 above allowed us to construct a non-dictatorial strategy-proof social choice function, but we will not be able to do so for all restricted domains; with other words, a restricted domain Ω either admits non-dictatorial strategy-proof social choice functions that satisfy unanimity, or not, and in the latter case, we will say that Ω is a *dictatorial domain*.

In the literature of strategy-proof social choice theory, restricted preference domains are studied from two different perspectives that complement each other: On the one hand, one is interested in to describe the structure of restricted domains that come up in a given context as a consequence of the nature of the available alternatives in combination with the economic incentives of the individuals in the society (see for instance Section 3.2). On the other hand, after having derived the structure of a certain restricted preference domain, one wants of course to know whether this domain admits non-dictatorial strategy-proof social choice functions, and therefore, a number of criteria has been developed to decide whether a given domain is dictatorial or non-dictatorial (see for instance Section 3.3). Note that the first two questions from the introduction combine these two aspects: Question 1 asks us to derive the structure of preferences over subsets of a fixed size, while Question 2 then asks whether this structure admits non-dictatorial strategy-proof social choice functions.

Finally, we would like to remark that a thorough introduction to restricted preference domains can be found in Sprumont (1995), and for a detailed survey of the most important restricted domains in social choice theory, we refer to the monograph of Gaertner (2001).

3.2 Single-peaked Preferences

The most common type of restricted preference domains is the class of *single-peaked preferences*, introduced by Black (1948) in order to model political preferences. It turns out that when the individuals in a society have single-peaked preferences, then the method of majority decision will lead to a transitive ordering of the available alternatives, which thus implies an escape from the impossibility result of Arrow's theorem. In strategy-proof social choice theory, single-peaked preferences are of similar interest because they admit non-dictatorial strategy-proof social choice functions, and it is therefore natural to ask whether the same is true for the social choice of fixed-sized subsets, whence we proposed Question 3 in the introduction. In the following, we illustrate how single-peaked preferences come up in applications, we present an appropriate formalization of single-peaked preferences, and we demonstrate the existence of a non-dictatorial strategy-proof social choice function on a domain of single-peaked preferences.

We begin with a political example. Suppose that the parties in an election can be ordered on a traditional left–right scale in the following way:

Strongly left party (SL) – *Modestly left party (ML)* – *Center party (C)* – *Modestly right party (MR)* – *Strongly right party (SR)*

If your most preferred party is *C*, then it seems reasonable that you will prefer *MR* to *SR*, and also *ML* to *SL*. Similarly, if your top alternative is *ML*, then you should prefer *C* to *MR* and *MR* to *SR*.

Consider next, as an economic example, possible preferences over the rate of inflation in a country. If you think that the optimal rate of inflation is two per cent, then you will probably prefer three per cent inflation to four per cent when the choice is between these two alternatives, and similarly, choosing between one per cent inflation and zero inflation, you prefer probably the former to the latter.

In both examples, the available alternatives can be ordered on a line, every individual can be assumed to have a most preferred alternative, the *peak*, and the more we get to the left respectively to the right of the peak, the less preferred are the alternatives. This is the structure that defines single-peaked preferences. We will now formalize this structure, and thereby, we follow closely the exposition in Mas-Colell et al. (1995). To begin with, the fact that alternatives can be ordered on a line will be modelled mathematically by a *linear order*:

Definition 3.1 (Linear Order). A *(strict) linear order* on a set $\mathscr{A}$ of alternatives is a binary relation $\prec$ on $\mathscr{A}$ that is complete, antisymmetric, and transitive.

If $\prec$ is a linear order on $\mathscr{A}$ and $a \prec b$, we say that a lies to the left of b, or equivalently, that b lies to the right of a. We will use the notation $a \preccurlyeq b$ to indicate that either $a \prec b$ or $a = b$. Sometimes, we also write $a \succ b$ (or $a \succcurlyeq b$) instead of $b \prec a$ (or $b \preccurlyeq a$). Starting with a linear order, a preference is single-peaked if alternatives are more desirable the closer they are to the most preferred alternative, or formally:

Definition 3.2 (Single-Peaked Preference). A preference P over the alternatives in $\mathscr{A}$ is said to be *single-peaked* with respect to the linear order $\prec$ if

$$\begin{aligned} & r_1(P) \preccurlyeq a \prec b \quad \Longrightarrow \quad aPb \\ \text{and} \qquad & r_1(P) \succcurlyeq a \succ b \quad \Longrightarrow \quad aPb. \end{aligned} \tag{3.2}$$

Preferences that satisfy (3.2) are called single-peaked because if the set of alternatives is a continuous variable, like the rate of inflation in the example above, then a single-peaked preference can be represented by utility function that is strictly increasing to the most preferred alternatives and then strictly decreasing, i.e., it has a single peak.

If a set $\mathscr{A}$ is equipped with a linear order $\prec$ and the number of individuals in the society is odd, then it is possible to define a social choice function that respects the underlying order of the alternatives. Consider first an example where a society has to choose one of four alternatives that are ordered according to $a_1 \prec a_2 \prec a_3 \prec a_4$, and suppose that a voting results in:

Alternative	a_1	a_2	a_3	a_4
Number of votes	3	1	1	2

Here, alternative a_2 can be claimed to be the natural outcome of this voting because it has equally many votes to its left as to its right. An alternative with this property is called a *median alternative*:

Definition 3.3 (Median Alternative). Suppose that the set $\mathscr{A}$ of alternatives is equipped with a linear order $\prec$, and let $\mathscr{P} \in \Sigma^N$ be a preference profile. An alternative $\bar{a}$ is said to be a *median alternative* for $\mathscr{P}$ if

$$\begin{aligned} & \sharp \left\{P_i \in \mathscr{P};\, r_1(P_i) \succcurlyeq \bar{a}\right\} \geq \frac{N}{2} \\ \text{and} \qquad & \sharp \left\{P_i \in \mathscr{P};\, r_1(P_i) \preccurlyeq \bar{a}\right\} \geq \frac{N}{2}. \end{aligned} \tag{3.3}$$

One can show that every preference profile has a median alternative, and in addition, when N is odd, then the median alternative turns out to be unique. In this case, one can thus define a social choice function f that assigns to each preference profile its median alternative, and this social choice function is known as the *median rule*. The median rule is of interest in this thesis because it is strategy-proof if preferences are single-peaked, a result which for instance can be found in Sprumont (1995) or Barberà (2001):

Theorem 3.1. *Let $\mathscr{A}$ be a set of alternatives equipped with a linear order $\prec$, and suppose that the number of individuals in the society is odd. If Ω is a domain of preferences over $\mathscr{A}$ that are single-peaked with respect to $\prec$, then the median rule*

$$f(\mathscr{P}) = \text{median}(\mathscr{P}) \quad \textit{for all } \mathscr{P} \in \Omega^N$$

is strategy-proof.

The median rule is of course non-dictatorial and satisfies unanimity, so Theorem 3.1 shows that strategy-proofness and non-dictatorship do not exclude each other when preferences are single-peaked. It is straight forward to argue why Theorem 3.1 must hold: Suppose that the median rule applied to the preference profile $\mathscr{P}$ leads to the social choice $\bar{a}$. If your own top alternative equals $\bar{a}$, then you can obviously not gain from misrepresentation. Consider therefore the case when your top alternative a differs from $\bar{a}$ and is located, say, to the right of $\bar{a}$, and suppose you are thinking of your possible gains from voting for some alternative a' instead of your true preference a. If a' also lies to the right of $\bar{a}$, or equals $\bar{a}$, then none of the cardinalities in (3.3) is affected, and your misrepresentation has thus no effect on the social choice. On the other hand, if a' lies to the left of $\bar{a}$, then either the median remains unaffected, or it is moved to the left, i.e., further away from your top alternative. Misrepresentation will therefore never be beneficial for you, and the median rule is thus strategy-proof.

In Chapter 5, we show that the median rule can be generalized to a social choice function for the social choice of fixed-sized subsets, and if preferences are single-peaked, then also this social choice function will turn out to be strategy-proof.

3.3 A Theoretical Tool: Linked Domains

The single-peaked preference domains considered in the previous section come up naturally in specific political or economic contexts, and they are easily accessible to economic intuition. On the contrary, the interest in linked domains, which we will discuss in this section, is not economicly motivated, but by the fact that linked domains provide a sufficient theoretical criterion for a domain to be dictatorial.

Linked domains have been introduced in Aswal et al. (2003), and Definition 3.4, Definition 3.5, and Theorem 3.2 in this section have been taken from this paper.

Recall from Section 3.1 that a restricted preference domain must contain sufficiently many preferences in order to be dictatorial. The criterion of linked domains shows that a domain is dictatorial if we can find sufficiently many inversions of alternatives in the top of the preferences. This will be made precise below, but first, we illustrate the underlying idea by an example: Let Ω be a preference domain over the set $\mathscr{A} = \{a, b, c\}$, and consider the following preferences, which are supposed to belong to Ω:

P_1	P_2	P_2'
a	b	b
b	a	c
c	c	a

Note that a and b appear in inverted order in the top of P_1 respectively P_2, and we will show that the existence of such preferences implies that every strategy-proof social choice function on Ω is at least partly dictatorial.[12] For simplicity we consider only the case when there are two individuals in the society, so suppose now that $f : \Omega^2 \to \mathscr{A}$ is a strategy-proof social choice function that satisfies unanimity. Our first observation is that we because of Pareto optimality (Lemma 2.5) either have $f(P_1, P_2) = a$ or $f(P_1, P_2) = b$, and we assume here that $f(P_1, P_2) = a$. Similarly, Pareto optimality implies also that $f(P_1, P_2') \in \{a, b\}$, but if $f(P_1, P_2') = b$, then individual 2 would be able to manipulate f by representing P_2' instead of P_2. Hence, we have $f(P_1, P_2') = a$, and monotonicity (Lemma 2.4) implies then that the social choice must be a whenever a is the top alternative of individual 1, and thus, individual 1 can be said to be a dictator for alternative a.

If there exist preferences P_1 and P_2 such that a and b are the two top alternatives in these preferences, but in inverted order, which was essential for the argument above, then we say that a and b are *connected*:

Definition 3.4 (Connectedness of Two Alternatives). Let $\mathscr{A}$ be a set of alternatives, and suppose that Ω is a restricted preferences domain over $\mathscr{A}$. Two alternatives $a_i, a_j \in \mathscr{A}$ are said to be *connected* in Ω if there exist both $P \in \Omega$ such that $r_1(P) = a_i$ and $r_2(P) = a_j$, and $P' \in \Omega$ such that $r_1(P') = a_j$ and $r_2(P') = a_i$. If a_i and a_j are connected, this will be denoted by $a_i \sim a_j$.[13]

[12] The following argument has been taken from Svensson (1999).

[13] Like the definition of connectedness of two alternatives, also the notation $a_i \sim a_j$ has been introduced by Aswal et al. (2003). There is no reason to deviate from it here, but we would like to point

For example, in the set Σ of unrestricted preferences, all pairs of alternatives are connected because every alternative can be ranked first and have any other alternative on the second rank. In this case, it is easy to extend the argument above in order to show that every strategy-proof social choice function $f : \Sigma^N \to \mathscr{A}$ that satisfies unanimity must be dictatorial, and it is actually in this way we prove the Gibbard-Satterthwaite theorem in Appendix B.

Before we turn to the criterion of linked domains, we show, for pedagogical purposes, how the notion of connectedness can be used to understand single-peaked preferences in a different way than in Section 3.2: If $a_1 \prec a_2 \prec \ldots \prec a_M$ is an underlying linear order of the elements in $\mathscr{A}$, then preferences that are single-peaked with respect to $\prec$ have the property that every a_i, with the exception of a_1, is connected to *exactly* one alternative in $\{a_1, a_2, \ldots, a_{i-1}\}$, namely to its direct predecessor a_{i-1}. We saw in the previous section that this structure is sufficiently restrictive to allow non-dictatorial strategy-proof social choice functions.

More generally, single-peaked preferences satisfy the following condition: The alternatives in $\mathscr{A}$ can be indexed in such way that every a_i is connected to at most one element in $\{a_1, a_2, \ldots, a_{i-1}\}$. Replacing *at most one* in this condition with *at least two* leads to much less restricted domains, called *linked domains*:

Definition 3.5 (Linked Domains). A preference domain Ω over $\mathscr{A}$ is said to be *linked* if the alternatives in $\mathscr{A}$ can be indexed in a sequence $a_1, a_2, \ldots, a_M$ in such a way that that $a_1 \sim a_2$ and every a_i with $i \geq 3$ is connected to at least two alternatives in $\{a_1, a_2, \ldots, a_{i-1}\}$.

Contrary to domains of single-peaked preferences, linked domains contain sufficiently many preferences to be dictatorial:

Theorem 3.2 (Theorem 3.1 in Aswal et al. (2003)). *Let $\mathscr{A}$ be a set of at least three alternatives, and suppose that Ω is a linked domain over $\mathscr{A}$. Then a social choice function $f : \Omega^N \to \mathscr{A}$ that satisfies unanimity is strategy-proof if and only if f is dictatorial.*

Theorem 3.2 should mainly be seen as a theoretical tool. For example, since the set Σ of unrestricted preference is obviously linked, the Gibbard-Satterthwaite theorem can be obtained as a direct consequence of Theorem 3.2. Note, however, that

out that the symbol $\sim$ is traditionally used in the economic literature to indicate that an individual is indifferent between two alternatives, see for instance Mas-Colell et al. (1995, 6). However, it should be clear that *connectedness* and *indifference* are two entirely different concepts.

the fact that a domain is linked is only a sufficient condition for a domain to be dictatorial, but it is not necessary, see Aswal et al. (2003, 46–47).

In this thesis, Theorem 3.2 plays the following role: If $\mathscr{A}$ is a set of alternatives, then the set of all preferences over subsets of $\mathscr{A}$ of a fixed size has a structure that is almost that of linked domains, but it consists of preferences that in general are not complete. In the following chapter, we will therefore generalize Theorem 3.2 to partial preference relations, and this result will then in Chapter 5 be used to conclude that non-dictatorial strategy-proof social choice of fixed-sized subsets is impossible in general.

4 Partial Preference Relations and Strategy-proof Social Choice

WE NOTED ALREADY in the introduction that preferences over subsets of a fixed size can fail to rank some subsets, and we will argue in Chapter 5 that these preferences are best modelled by partial preference relations. Therefore, we present in Section 4.1 in this chapter a short general introduction to partial preferences and we introduce some notions to describe their structure. In Section 4.2, we adjust the notion of linked domains in such a way that it can be applied to preferences over fixed-sized subsets, and we will generalize Theorem 3.2 to a large class of partial preferences. In Section 4.3, finally, we answer, as a first application of our generalization, Question 4 from the introduction.

4.1 Partial Preference Relations

Partial preference relations are preferences that are antisymmetric and transitive, but contrary to the preferences considered in the previous two chapters, they are not necessarily complete, which means that a partial preference may fail to rank some pairs of alternatives. Partial preferences are thus more general than complete preferences, and every result that holds for partial preferences holds also for complete preferences. In the economic literature, however, the use of partial preferences is quite limited because it seems natural to assume that an individual that is facing two alternatives a and b either prefers a to b, or b to a, or regards them as equally good. To involve a fourth possible attitude, namely, that the individual does not rank a and b at all, seems to be a needless complication. In general, however, there are at least two aspects that can motivate the use of partial preference relations.

Firstly, the nature of available alternatives can be such that they are not unambiguously comparable, for example, because they have several independent quality dimensions that can come into conflict. In essence, this conflict is about how to compare a comfortable house at an unattractive place with a less comfortable house at an attractive place. Confronted with such a choice situation, many people feel uncomfortable and might be unwilling to express any preference. It might be tempting, at first glance, to consider this inability or unwillingness to rank two alternatives as being the same as indifference. However, in Section 5.1, we will show that this view leads to a contradiction (see page 50), which means that the absence of explicit preference or indifference as it is modelled by partial preference relations differs not only philosophically from complete preferences, but can also lead to different, possibly more reasonable, results in the mathematical formalization.

Secondly, facing a number of alternatives people are often truly engaged in only one or two of them and prefer these to any other alternative, but they are not especially interested in ranking the remaining alternatives internally. This kind of preference structure, which is particularly likely when the number of alternatives is large, is conveniently modelled by partial preferences.

Beside these general aspects, there are economic applications that give rise to a more direct need for partial preference relations because the available information about individual preferences may be insufficient. This is, for instance, the case for the voting problem considered in this thesis because in Chapter 5, we will show that if voters have complete preferences over the alternatives in a set $\mathscr{A}$, then these preferences can be used to rank some subsets of $\mathscr{A}$, whereas other subsets cannot be ranked with this information.

In many cases in economics, it may thus be a good choice to model individual preferences by partial preference relations. Often, however, it will be convenient to assume that even incomplete preferences have more structure than only antisymmetry and transitivity, and we introduce now two types of partial preferences with additional structure. As pointed out above, it is often reasonable to assume that an individual has a special interest in some of the available alternatives, and the weakest economicly meaningful assumption is that an individual at least can point out one of the alternatives to be his most preferred. Preferences with this property will be called *top-1 preference relations*:

Definition 4.1 (Top-1 Preference Relation). A partial preference relation P on a set $\mathscr{A}$ of alternatives is said to be a *top-1 preference relation* on $\mathscr{A}$, if there exists an alternative $a \in \mathscr{A}$ such that

$$a\,P\,x \quad \text{for all} \quad x \in \mathscr{A} \setminus \{a\}. \tag{4.1}$$

A set of top-1 preference relations on $\mathscr{A}$ will be called a *top-1 domain*.

If P is a top-1 preference relation, then the alternative $a \in \mathscr{A}$ satisfying (4.1) is unique,[14] and it will henceforth be denoted by $r_1(P)$. Note also, that condition (4.1) is only a minimal requirement on a top-1 preference and does not exclude that P also ranks other alternatives. Starting with the concept of top-1 preference relations

[14]This follows easily from the antisymmetry property of a partial preference: Suppose there are two different alternatives $a, a' \in \mathscr{A}$ satisfying (4.1). Then we must have both $a\,P\,a'$ and $a'\,P\,a$, but as this contradicts antisymmetry, we conclude that there can be at most one $a \in \mathscr{A}$ satisfying (4.1).

it is straight forward to add more structure by requiring not only a top alternative, but also a second best alternative:

Definition 4.2 (Top-2 Preference Relation). A top-1 preference relation P on a set $\mathscr{A}$ is said to be also a *top-2 preference relation* on $\mathscr{A}$, if there exists an alternative $a \in \mathscr{A} \setminus \{r_1(P)\}$ such that

$$a\,P\,x \quad \text{for all} \quad x \in \mathscr{A} \setminus \{r_1(P), a\}. \tag{4.2}$$

A set of top-2 preference relations on $\mathscr{A}$ will be called a *top-2 domain*, and the unique element a satisfying (4.2) will be denoted by $r_2(P)$.

Top-2 preferences are introduced in this thesis because preferences over subsets of a fixed size turn out to have precisely this structure, which we will see in Chapter 5. But in general, top-2 preferences can also be useful in other economic applications, which the following example illustrates.

Example 4.1. Students at Lund university applying for spring term 2006 had the possibility to choose among 208 different beginners' courses. On the application form, the students were allowed to fill in a main alternative and a reserve alternative. Most students know that this can be a sufficiently demanding task, and it is not very reasonable to assume that students have complete preferences over all courses. Hence, it could be a good idea to model students' preferences over courses by top-2 preference relations. □

Remark 4.1. After having defined top-2 preference relations from top-1 preference relations, one can of course continue to add more structure to preference relations by requiring that an individual is able to report a top of his preferences consisting of three, four, or generally k top alternatives, which would lead us to a definition of *top-k preference relations*. However, since this more general concept is not needed in the sequel, we refrain from considering it in detail. □

Note finally that a partial preference can be thought of as a part of complete preference in the following sense: A complete preference $\bar{P}$ is said to be *compatible* with the partial preference P if any ranking of two alternatives that holds under P also holds under $\bar{P}$, that is, if

$$a\,P\,b \quad \Longrightarrow \quad a\,\bar{P}\,b$$

for all $a, b \in \mathscr{A}$. Given a partial preference P, there exists always a complete preference compatible with P, and in general there will actually be several such complete preferences. This result, which is known as *Szpilrajn's theorem*, has been proved in Szpilrajn (1930), and it will be used in Chapter 5.

4.2 Strategy-proof Social Choice on Linked Top-2 Domains

In this section, we generalize Theorem 3.2 to top-2 preference relations. To begin with, we re-formulate in Section 4.2.1 the definitions that are used in Theorem 3.2 for top-2 domains, and we state our generalization of Theorem 3.2. This result is one of the main contributions of this thesis, and we will therefore prove it in detail. Section 4.2.2 introduces some preparatory results that are frequently used in the proof, and Section 4.2.3 contains the complete proof.

4.2.1 Basic Definitions and Statement of the Main Theorem

All notions used in the formulation of Theorem 3.2 have been defined under the silent assumption that preferences are complete. In the following, we check therefore that these definitions can be transferred to top-2 preference relations, and we introduce also a new notion of manipulability, which is more appropriate when preferences belong to a top-2 domain.

The basic scenario is now as follows: A society consisting of N individuals, indexed by the set $\mathcal{I} = \{1, 2, \ldots, N\}$, must choose one element from a set $\mathscr{A}$ that contains M alternatives. The individuals are assumed to have top-2 preferences over the alternatives in $\mathscr{A}$, and the set of all admissible preferences will be denoted by the Greek letter Γ.[15] The social choice is made using a social choice function which now has the form $f : \Gamma^N \to \mathscr{A}$.

The definitions of dictatorship, respect of unanimity, and linked domains can now be applied almost without modifications, but in order to avoid any confusions, we will re-formulate these definitions explicitly: A social choice function $f : \Gamma^N \to \mathscr{A}$ is said to be *dictatorial* if there exists an individual i, the *dictator*, such that $f(P_1, P_2, \ldots, P_N) = r_1(P_i)$ for all preferences profiles $(P_1, P_2, \ldots, P_N) \in \Gamma^N$. Next, we will say that $f : \Gamma^N \to \mathscr{A}$ satisfies *unanimity with respect to alternative* $a \in \mathscr{A}$ if $f(P_1, P_2, \ldots, P_N) = a$ for all preference profiles $(P_1, P_2, \ldots, P_N) \in \Gamma^N$ such that $r_1(P_i) = a$ for every $i \in \mathcal{I}$, and if f satisfies unanimity with respect to every alternative in $\mathscr{A}$, we simply say that f satisfies *unanimity*. Note that these definitions can be transferred to top-2 preferences because the top alternative $r_1(P)$ is well-defined for every top-2 preference P. Since top-2 preferences also have a second best alternative, it is also possible to translate the definition of linked domains. To begin with, we will say that two alternatives $a_i, a_j \in \mathscr{A}$ are *connected* in a top-2 domain Γ if there exist both $P \in \Gamma$ such that $r_1(P) = a_i$ and $r_2(P) = a_j$, and $P' \in \Gamma$ such

[15] Γ is not necessarily the set of *all* top-2 preferences over $\mathscr{A}$, but Γ can be any top-2 domain.

that $r_1(P') = a_j$ and $r_2(P') = a_i$, and if a_i and a_j are connected, this will as before be denoted by $a_i \sim a_j$. A top-2 domain Γ is said to be *linked* if the alternatives in $\mathscr{A}$ can be indexed as a sequence $a_1, a_2, \ldots, a_M$ in such a way that $a_1 \sim a_2$ and every a_i with $i \geq 3$ is connected with at least two alternatives in $\{a_1, a_2, \ldots, a_{i-1}\}$. For later use, we introduce also one more notation: If Γ is a linked top-2 domain and $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ are n alternatives in $\mathscr{A}$, we will use the short form $[a_{i_1}, a_{i_2}, \ldots, a_{i_n}]$ to indicate that

$$a_{i_1} \sim a_{i_2}\,, \quad a_{i_2} \sim a_{i_3}\,, \quad \ldots, \quad a_{i_{n-2}} \sim a_{i_{n-1}}\,, \quad \text{and} \quad a_{i_{n-1}} \sim a_{i_n},$$

and we will say that $[a_{i_1}, a_{i_2}, \ldots, a_{i_n}]$ is a *chain* connecting a_{i_1} with a_{i_n}. Note that if a_i and a_j are two alternatives in $\mathscr{A}$ that are not connected, then it is always possible to find a chain that connects a_i and a_j.[16]

We turn now to the notions of manipulability and strategy-proofness for top-2 preferences. Of course, a social choice function $f : \Gamma^N \to \mathscr{A}$ can, precisely as in the case of complete preferences, be defined to be manipulable if some individual has incentives to misrepresent his preferences:

Definition 4.3 (Manipulability and Strategy-proofness). A social choice function $f : \Gamma^N \to \mathscr{A}$ is said to be *manipulable* if there exist preferences $P_i, P_i' \in \Gamma$ and some preference profile $P_{-i} \in \Gamma^{N-1}$ such that

$$f(P_i', P_{-i}) \;\; P_i \;\; f(P_i, P_{-i}). \tag{4.3}$$

If f is not manipulable, we say that f is *strategy-proof.*

Note that (4.3) is precisely the same condition as before, but since P_i now only is assumed to be a partial preference, which in general consists of fewer rankings than a complete preference does, it should be more difficult to find a P_i' such that (4.3) is satisfied, and therefore, it should a priori be easier to find strategy-proof social choice functions. For reasons to be explained below we will here in connection with top-2 preferences not use the definition above, but we will work with a more informative notion of manipulability that only focuses on whether some individual can obtain one of his two top alternatives:

---

[16]This can be seen in the following way: If $a_i \neq a_1$, then there exists an alternative preceding a_i in $a_1, a_2, \ldots, a_M$ that is connected with a_i, and this alternative has also a preceding alternative to which it is connected, and so on. Thus, there must be a chain connecting a_i with a_1. Similarly, there exists also a chain connecting a_1 with a_j, and joining the two chains gives a chain that connects a_i with a_j.

Definition 4.4 (Top-2 Manipulability and Top-2 Strategy-proofness). If Γ is a top-2 domain, we say that a social choice function $f : \Gamma^N \to \mathscr{A}$ is *top-2 manipulable* if there exist preferences $P_i, P_i' \in \Gamma$ and some preference profile $P_{-i} \in \Gamma^{N-1}$ such that

$$f(P_i', P_{-i}) \; P_i \; f(P_i, P_{-i}) \quad \text{and} \quad f(P_i', P_{-i}) = r_1(P_i) \quad \text{or} \quad f(P_i', P_{-i}) = r_2(P_i).$$

If f is not top-2 manipulable, we say that f is *top-2 strategy-proof.*

Definition 4.3 and Definition 4.4 are of course related in the following way: A strategy-proof social choice function is also top-2 strategy-proof, and a top-2 manipulable social choice function is also manipulable, but the converse implications need not be true. The notion of top-2 manipulability is introduced here for two reasons: Firstly, out generalization of Theorem 3.2 shows that a large class of social choice functions actually is top-2 manipulable, and it is thus natural to use this notion of manipulability because it is more informative than ordinary manipulability. Secondly, some steps in the proof of our generalization of Theorem 3.2 are only valid if we assume top-2 manipulability, which forces us to work with this notion of manipulability.[17]

We are now able to state our generalization of Theorem 3.2, which will be proved in the remainder of this section:

Theorem 4.1. *Let Γ be a linked top-2 domain over a finite set $\mathscr{A}$, and assume that $f : \Gamma^N \to \mathscr{A}$ is a social choice function that satisfies unanimity. Then f is top-2 strategy-proof if and only if f is dictatorial.*

4.2.2 Some Preparatory Results

In this subsection, we derive three basic result which will be needed in the formal proof of Theorem 4.1 in the next subsection. In the proof of the original Gibbard-Satterthwaite theorem in Appendix B, the properties of monotonicity and Pareto optimality, which we already considered in Chapter 2, play an important role, and

[17] Beside these concrete reasons, there are also economic motivations for the use of top-2 strategy-proofness: As pointed out above, people facing a number of alternatives are often truly engaged in only a few of them. If it is reasonable to assume that individuals only are engaged in two alternatives, preferring these to all other alternatives without ranking any other pair of alternatives, then no individual will have incentives to misrepresent his preferences when a top-2 strategy-proof social choice function is used, which means that top-2 strategy-proofness in practice implies strategy-proofness. On the other hand, if a social choice function is top-2 manipulable, then some individual can at some instance by misrepresentation obtain an alternative in which he his truly engaged, and therefore individuals have incentives to think about tactical voting.

they will also be central in our proof of Theorem 4.1. The definitions of monotonicity and Pareto optimality given in Chapter 2 are however only applicable to complete preferences, and therefore, we introduce here weaker variants of these properties, which then are shown to be valid for top-2 strategy-proof social choice functions. We will also prove a reduction lemma, which is needed in the proof of Theorem 4.1.

The monotonicity property of strategy-proof social choice functions will be weakened in the following way: If P and P' are two preferences in a top-2 domain Γ, we will say that an alternative $a \in \mathscr{A}$ *moves to the top* from P to P' whenever $r_1(P') = a$, that is, even if a already is the top alternative of P, and we have then the following lemma:

Lemma 4.2 (Monotonicity for Top-moving Alternatives). *Let Γ be a top-2 domain, and suppose that $f : \Gamma^N \to \mathscr{A}$ is a top-2 strategy-proof social choice function. If the preference profile $(P_i, P_{-i}) \in \Gamma^N$ gives $f(P_i, P_{-i}) = a$, and a moves to the top from P_i to the preference $P_i' \in \Gamma$, then we also have $f(P_i', P_{-i}) = a$.*

Proof. We argue by contradiction and assume that $f(P_i', P_{-i}) = b$ and $b \neq a$. Since $r_1(P_i') = a$, we have $a P_i' b$, and individual i can manipulate f by going from (P_i', P_{-i}) to (P_i, P_{-i}). As this contradicts the top-2 strategy-proofness of f, the assumption $f(P_i', P_{-i}) \neq a$ must have been wrong, and the lemma is proved. □

Note that monotonicity for top-moving alternatives of course is a special case of the monotonicity property from Chapter 2. We will frequently use a generalized monotonicity property of top-2 strategy-proof social choice functions, which easily follows from Lemma 4.2: Suppose that a preference profile $(P_1, P_2, \ldots, P_N) \in \Gamma^N$ gives $f(P_1, P_2, \ldots, P_N) = a$. If $(P_1', P_2', \ldots, P_N') \in \Gamma^N$ is another preference profile such that either $r_1(P_i') = a$ or $P_i' = P_i$ for all $i \in \mathcal{I}$, that is, either a moves to the top of an individual's preference, or an individual's preference remains unchanged, then repeated use of Lemma 4.2 shows that also $f(P_1', P_2', \ldots, P_N') = a$.

Also in the case of Pareto optimality, we are only able to prove a much weaker variant of the Pareto optimality considered in Chapter 2.

Lemma 4.3 (Simple Pareto Optimality). *Let Γ be a top-2 domain, and suppose that $f : \Gamma^N \to \mathscr{A}$ is a top-2 strategy-proof social choice function. Let further a and b be two distinct alternatives in $\mathscr{A}$, and suppose that f satisfies unanimity with respect to a. If $\mathscr{P} \in \Gamma^N$ is a preference profile such that every $P_i \in \mathscr{P}$ satisfies either $r_1(P_i) = a$, or $r_1(P_i) = b$ and $r_2(P_i) = a$, then either $f(\mathscr{P}) = a$ or $f(\mathscr{P}) = b$.*

Proof. Note first that if all $P_i \in \mathscr{P}$ satisfy $r_1(P_i) = a$, then $f(\mathscr{P}) = a$ by unanimity. Next, consider the case when exactly one individual, say individual 1 for simplicity, ranks b first, that is, $r_1(P_1) = b$, $r_2(P_1) = a$, and $r_1(P_i) = a$ for all $i \geq 2$. If we would have $f(P_1, P_{-1}) = c$ for some $c \in \mathscr{A} \setminus \{a, b\}$, then individual 1 would be able to top-2 manipulate f because $f(P_1', P_{-1}) = a$ for every $P_1' \in \Gamma$ with $r_1(P_1') = a$, and hence, we conclude that $f(P_1, P_{-1}) \in \{a, b\}$. If exactly two individuals, say individual 1 and individual 2, rank b first and $f(P_1, P_2, \ldots, P_N) = c$ for some $c \in \mathscr{A} \setminus \{a, b\}$, then individual 2 is able to top-2 manipulate f because by the previous argument $f(P_1, P_2', P_3, \ldots, P_N) \in \{a, b\}$ for every $P_2' \in \Gamma$ with $r_1(P_2') = a$, and we can thus conclude that $f(P_1, P_2, \ldots, P_N) \in \{a, b\}$. The lemma follows now when we successively consider the cases when exactly three, four, and so forth, individuals rank b first. □

In the proof of Theorem 4.1, we will repeated times consider preference profiles where all but one individual in the society have the same preference, and thereby, we will need the following reduction lemma.

Lemma 4.4 (Reduction Lemma). *Let Γ be a top-2 domain, and suppose that $f : \Gamma^N \to \mathscr{A}$ is a top-2 strategy-proof social choice function. Construct, for some fixed $i \in \mathcal{I}$, the social choice function $f_i : \Gamma^2 \to \mathscr{A}$ from f by setting all arguments in f equal to P_2, except from the* ith *one, where we insert P_1, that is*

$$f_i(P_1, P_2) = f(P_2, \ldots, \underset{i-1}{P_2}, \underset{i}{P_1}, \underset{i+1}{P_2}, \ldots, P_2). \tag{4.4}$$

If f is top-2 strategy-proof, then also f_i is top-2 strategy-proof.

Proof. Suppose, for simplicity, that $i = 1$. Since f is top-2 strategy-proof, it is clear that f_1 is top-2 strategy-proof with respect to P_1. To show that f_1 also is top-2 strategy-proof with respect to P_2, we will argue by contradiction. Suppose therefore that there exist $a, b \in \mathscr{A}$ and $P_1, P_2, P_2' \in \Gamma$ such that

$$f_1(P_1, P_2) = b, \quad f_1(P_1, P_2') = a, \quad a P_2 b, \quad \text{and} \quad a \in \{r_1(P_2), r_2(P_2)\}. \tag{4.5}$$

By the definition of f_1, we have

$$\begin{aligned} & f_1(P_1, P_2) = b \iff f(P_1, P_2, P_2, \ldots, P_2) = b, \\ \text{and} \quad & f_1(P_1, P_2') = a \iff f(P_1, P_2', P_2', \ldots, P_2') = a. \end{aligned}$$

Changing the preferences in the argument of f successively from $(P_1, P_2, \ldots, P_2)$ to $(P_1, P_2', \ldots, P_2')$ for one individual at a time, that is, first for individual 1, then for individual 2, and so forth, we conclude that there must be an instance such that

$$f(P_1, P'_2, \ldots, P'_2, P_2, P_2, \ldots, P_2) = c, \tag{4.6}$$

for some $c \in \mathscr{A} \setminus \{a\}$, but

$$f(P_1, P'_2, \ldots, P'_2, P'_2, P_2, \ldots, P_2) = a. \tag{4.7}$$

We claim that $a P_2 c$. This will prove the lemma because (4.6) and (4.7) combined with $a P_2 c$ and $a \in \{r_1(P_2), r_2(P_2)\}$ contradict the top-2 strategy-proofness of f. We have to consider two cases, firstly, $a = r_1(P_2)$, and secondly, $a = r_2(P_2)$. If $a = r_1(P_2)$, it is obvious that $a P_2 c$. On the other hand, if $a = r_2(P_2)$, we must show that $c \neq r_1(P_2)$. We argue by contradiction and assume that $r_1(P_2) = c$. Applying monotonicity of top-moving alternatives to (4.6), we obtain $f(P_1, P_2, \ldots, P_2) = c$, and since also $f(P_1, P_2, \ldots, P_2) = b$, we conclude that $b = c$. But then $b = c = r_1(P_2)$ implies $b P_2 a$, which contradicts the assumption $a P_2 b$ in (4.5). Thus $c \neq r_1(P_2)$, and we are done. □

4.2.3 The Proof of Theorem 4.1

We turn now to the proof of Theorem 4.1, the interesting part of which of course is to show that top-2 strategy-proofness implies dictatorship. This will be proved in the following by a chain of lemmas, and in order to avoid repeated formulations, we assume now that the following assumptions hold throughout this subsection: We assume that Γ is a linked top-2 domain over a set $\mathscr{A}$, which contains $M \geq 3$ alternatives, and we suppose that $a_1, a_2, \ldots, a_M$ is an indexing of the alternatives in $\mathscr{A}$ that satisfies (1) $a_1 \sim a_2$ and (2) every a_i with $i \geq 3$ is connected to at least two alternatives in $\{a_1, a_2, \ldots, a_{i-1}\}$. Further, we consider a social choice function $f : \Gamma^N \to \mathscr{A}$ that is assumed to be top-2 strategy-proof.

For a clearer exposition, we will at several places present top-2 preferences in tables where we indicate the two top alternatives; for example, we write

P_i	P_{-i}
a_1	a_3
a_2	$\vdots$

in order to indicate that $P_i \in \Gamma$ is a preference such that $r_1(P_i) = a_1$ and $r_2(P_i) = a_2$, whereas every preference P_j in the preference profile P_{-i} satisfies $r_1(P_j) = a_3$.

The proof of the fact that top-2 strategy-proofness implies dictatorship consists essentially of two parts. In the first part (Lemma 4.5 to Lemma 4.9), we investigate how the preference of a single individual affects the set of possible social choices, that is, we ask what alternatives are left for society when individual i has reported his preference. Thereby, the set of all social choices that still are available when individual i has chosen preference P_i will be denoted by

$$\mathscr{O}_{-i}(P_i, f) = \{a \in \mathscr{A};\ a = f(P_i, P_{-i}) \text{ for some } P_{-i} \in \Gamma^{N-1}\},$$

and it is called the *option set* with respect to P_i and f. When the social choice function f is obvious from the context, we will shortly write $\mathscr{O}_{-i}(P_i)$ instead of $\mathscr{O}_{-i}(P_i, f)$. The idea behind this approach is that if there exists a dictator for f, which we want to prove, then there should be some $i \in \mathcal{I}$ such that society's possible choices are always restricted to individual i's top alternative, that is, $\mathscr{O}_{-i}(P_i)$ should collapse to $\{r_1(P_i)\}$ for all $P_i \in \Gamma$. To begin with, we will in two respects restrict our attention to the first three elements in $\mathscr{A}$ only: firstly, we will only consider preferences whose top alternative is one of a_1, a_2, or a_3, and secondly, we will only investigate whether the preference of a single individual affects society's possibility to choose among a_1, a_2, or a_3. Formally, this means we will analyze the set

$$\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} \qquad \text{for} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3,$$

where Γ_k for $k \in \{1, 2, \ldots, M\}$ denotes the set of all preferences in Γ that have a_k as top alternative. At the end of the first part of the proof, we will have shown that there exists exactly one individual i such that f, under a weak additional condition, always chooses individual i's top alternative if this is one of a_1, a_2, or a_3. This individual should then clearly be regarded as a seed for a dictator for f, and in the second part of the proof, we will show that f still chooses individual i's top alternative when we successively extend the set of admissible preferences, first for the other individuals in the society (Lemma 4.10), and then also for the presumed dictator (Lemma 4.11). In a final step, we remove the weak restriction mentioned above and show that individual i indeed is a dictator for f.

Throughout the proof, two arguments will be used repeated times. For the sake of clarity, they are presented here, and when we use them in the sequel, we will only shortly refer to them as *Observation 1* respectively *Observation 2*.

Observation 1: If $\mathscr{P} = (P_1, P_2, \ldots, P_N)$ is a preference profile such that $a \in \mathscr{O}_{-i}(P_i)$ for some $i \in \mathcal{I}$, and $r_1(P_j) = a$ for all P_j with $j \neq i$, then $f(P_1, P_2, \ldots, P_N) = a$. This can be seen in the following way: Since $a \in \mathscr{O}_{-i}(P_i)$, there must be a preference profile $\mathscr{P}' = (P_i, P'_{-i})$ such that $f(P_i, P'_{-i}) = a$. Going from P'_{-i} to P_{-i}, alternative a moves to the top of all preferences in P_{-i}, and since f is assumed to be strategy-proof, we can apply Lemma 4.2 in order to conclude that $f(P_1, P_2, \ldots, P_N) = a$. $\square$

Observation 2: Let a and b be two distinct alternatives in $\mathscr{A}$, and suppose that $\mathscr{P} = (P_1, P_2, \dots, P_N)$ is a preference profile such that $r_1(P_i) = a$ and $r_2(P_i) = b$ for some $i \in \mathcal{I}$, and for all $j \in \mathcal{I} \setminus \{i\}$, we have $r_1(P_j) = b$, that is

P_1	$\cdots$	P_{i-1}	P_i	P_{i+1}	$\cdots$	P_N
b	$\cdots$	b	a	b	$\cdots$	b
$\vdots$	$\cdots$	$\vdots$	b	$\vdots$	$\cdots$	$\vdots$

If $b \notin \mathscr{O}_{-i}(P_i)$ and f satisfies unanimity with respect to a, then $f(\mathscr{P}) = a$. This follows easily once we noted that simple Pareto optimality implies $f(\mathscr{P}) \in \{a, b\}$, because then $b \notin \mathscr{O}_{-i}(P_i)$ excludes the case $f(\mathscr{P}) = b$, and hence $f(\mathscr{P}) = a$. Similarly, if $r_1(P_i) = a$ and $r_1(P_j) = b$ and $r_2(P_j) = a$ for every $P_j \in P_{-i}$, then $b \notin \mathscr{O}_{-i}(P_i)$ implies also $f(\mathscr{P}) = a$, provided that f satisfies unanimity with respect to a. $\square$

We enter now the first part of the proof. As indicated before, our attention will here be restricted to a_1, a_2, and a_3, and therefore we will from Lemma 4.5 to Lemma 4.9 only require that f satisfies unanimity with respect to a_1, a_2, and a_3 (in the second part, however, we will require that f satisfies unanimity with respect to all alternatives in $\mathscr{A}$). The first three lemmas in the proof (Lemma 4.5 to Lemma 4.7) clarify the structure of the set $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\}$. Thereby, it will turn out that an individual i either can reduce the set of possible social choices to his top alternative, or his preference P_i has no impact on this set. The first lemma shows that for a fixed top alternative, say for simplicity $r_1(P_i) = a_1$, the set $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\}$ does not depend on the particular choice of $P_i \in \Gamma_1$.

Lemma 4.5. *Assume that $P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$. Then*

$$\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \mathscr{O}_{-i}(\bar{P}_i) \cap \{a_1, a_2, a_3\}$$

for all $\bar{P}_i$ with $r_1(\bar{P}_i) = r_1(P_i)$.

Proof. Assume, without loss of generality, that $r_1(P_i) = r_1(\bar{P}_i) = a_1$. The lemma will be proved by a contradiction argument. Assume therefore further, again without loss of generality, that $a_2 \in \mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\}$, but $a_2 \notin \mathscr{O}_{-i}(\bar{P}_i) \cap \{a_1, a_2, a_3\}$. Since $a_1 \sim a_2$, we can choose some preference profile P_{-i} such that $r_1(P_j) = a_2$ and $r_2(P_j) = a_1$ for all $j \in \mathcal{I} \setminus \{i\}$. On the one hand, since $a_2 \in \mathscr{O}_{-i}(P_i)$, we have $f(P_i, P_{-i}) = a_2$ by Observation 1. On the other hand, $f(\bar{P}_i, P_{-i}) = a_1$ according to Observation 2. But since $r_1(P_1) = a_1$, this means that individual i can top-2 manipulate f by going from (P_i, P_{-i}) to $(\bar{P}_i, P_{-i})$, which contradicts the top-2 strategy-proofness of f. The lemma is proved. $\square$

Next, we will see that $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\}$ either contains the top alternative in P_i or all of a_1, a_2, and a_3. Note that the former case implies that individual i can reduce society's possible choice drastically, whereas the latter case means the he has no such power.

Lemma 4.6. *Assume that* $P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$. *Then either*

$$\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\} \quad \textit{or} \quad \mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\}.$$

Proof. Again, we will argue by contradiction. Suppose therefore, without loss of generality, that $P_i \in \Gamma_1$ and $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{a_1, a_2\}$. Moreover, in spite of the previous lemma, we can also assume that $r_2(P_i) = a_3$. Consider now preference profiles P_{-i} and P'_{-i} such that

P_i	P_{-i}	P'_{-i}
a_1	a_3	a_2
a_3	a_2	a_3

By Observation 2, we have $f(P_i, P_{-i}) = a_1$. On the other hand, from Observation 1 follows $f(P_i, P'_{-i}) = a_2$. But this means that the function f_i defined by (4.4) is top-2 manipulable, which by Lemma 4.4 contradicts the top-2 strategy-proofness of f, and the lemma is proved. ☐

An individual that can reduce society's possible choices with respect to one top alternative is in fact able to do this with respect to all of a_1, a_2, and a_3:

Lemma 4.7. *For every* $i \in \mathcal{I}$, *we have either*

$$\begin{aligned} & \mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\} \qquad \textit{for all} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \\ \textit{or} \quad & \mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\} \qquad \textit{for all} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3. \end{aligned}$$

Proof. If $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\}$ for all $P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, the lemma is obviously true. Consider therefore the case when $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\}$ for some $P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, and we assume without loss of generality that $P_i \in \Gamma_1$. Moreover, due to Lemma 4.5, we can also assume that $r_2(P_i) = a_3$. Note that such a P_i exists because $a_1 \sim a_3$. Suppose now, contrary to the claim of the lemma, that for some $P'_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ we have $\mathscr{O}_{-i}(P'_i) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\}$. From Lemma 4.5 it is clear that $P'_i \notin \Gamma_1$; we assume without loss of generality that $P'_i \in \Gamma_2$, and by Lemma 4.5, we can also assume that $r_2(P'_i) = a_1$. Consider now

P_i	P'_i	P_{-i}
a_1	a_2	a_3
$\vdots$	a_1	a_1

As a consequence of Observation 2, we must have $f(P_i, P_{-i}) = a_1$. On the other hand, Observation 1 implies $f(P_i', P_{-i}) = a_3$. But this means that individual i can manipulate f from (P_i', P_{-i}) to (P_i, P_{-i}). Thus, if $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\}$ for $P_i \in \Gamma_1$, then $\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\}$ for all $P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, and the lemma is proved. □

Up to this point, we have shown that an individual either can restrict society's choice to his favour, or that he has no such power, without having ensured the existence of either of these types. Next, we show that there in deed exists one individual whose choice restricts society's possible choices, and this individual must then of course be unique.

Lemma 4.8. *There is exactly one $i \in \mathcal{I}$ such that*

$$\mathscr{O}_{-i}(P_i) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\} \qquad \textit{for all} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3. \tag{4.8}$$

Proof. Note first that there can be at most one $i \in \mathcal{I}$ satisfying (4.8). To see this suppose that (4.8) holds for at least two individuals, and assume for simplicity that these are individual 1 and 2. Thus

$$\mathscr{O}_{-1}(P_1) \cap \{a_1, a_2, a_3\} = \{r_1(P_1)\} \quad \text{for all} \quad P_1 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3, \tag{4.9}$$

$$\text{and} \qquad \mathscr{O}_{-2}(P_2) \cap \{a_1, a_2, a_3\} = \{r_1(P_2)\} \quad \text{for all} \quad P_2 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3. \tag{4.10}$$

Consider now a preference profile $\mathscr{P} = (P_1, P_2, \ldots, P_N)$ where $r_1(P_j) = a_1$ and $r_2(P_j) = a_2$ for all $j \geq 3$, and P_1 and P_2 are such that

P_1	P_2	P_3	$\cdots$	P_N
a_1	a_2	a_1	$\cdots$	a_1
a_2	a_1	a_2	$\cdots$	a_2

Simple Pareto optimality implies $f(\mathscr{P}) \in \{a_1, a_2\}$, and since $a_2 \notin \mathscr{O}_{-1}(P_1)$ by (4.9), we must have $f(\mathscr{P}) = a_1$. But similarly, it follows from (4.10) that $f(\mathscr{P}) = a_2$. This contradiction shows that at most one $i \in \mathcal{I}$ can satisfy (4.8).

The other part of the lemma, namely that there exists at least one $i \in \mathcal{I}$ for which (4.8) holds, will be proved by induction over the number of individuals. Consider thus first the case when $N = 2$, and suppose, contrary to the claim of the lemma, that neither individual 1 nor 2 satisfies (4.8). In the light of the previous lemma, this means

$$\mathscr{O}_{-1}(P_1) \cap \{a_1, a_2, a_3\} = \mathscr{O}_{-2}(P_2) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\}$$

for all $P_1, P_2 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$. Consider a preference profile (P_1, P_2) with $r_1(P_1) = a_1$ and $r_1(P_2) = a_2$. Since $a_1 \in \mathscr{O}_{-2}(P_2)$ and $a_1 = r_1(P_1)$, we have $f(P_1, P_2) = a_1$ according to Observation 1. But similarly, since $a_2 \in \mathscr{O}_{-1}(P_1)$ and $a_2 = r_1(P_2)$, it follows also that $f(P_1, P_2) = a_2$. This, however, contradicts the assumption that f is a well-defined function, and the basis step of the induction is established.

For the induction step, suppose now that the claim is true for $N = n$, and assume that $f : \Gamma^{n+1} \to \mathscr{A}$ is defined for a society of size $n+1$. For individual $n+1$, we have according to the previous lemma either

$$\begin{aligned} & \mathscr{O}_{-(n+1)}(P_{n+1}) \cap \{a_1, a_2, a_3\} = \{r_1(P_{n+1})\} \quad \text{for all} \quad P_{n+1} \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3, \\ \text{or} \quad & \mathscr{O}_{-(n+1)}(P_{n+1}) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\} \quad \text{for all} \quad P_{n+1} \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3. \end{aligned}$$

In the first case, the claim follows at once, and we assume therefore in the continuation that the second case holds. Fix some $\bar{P}_{n+1}$ with $r_1(\bar{P}_{n+1}) = a_1$, and consider the function $\tilde{f} : \Gamma^n \to \mathscr{A}$ defined by

$$\tilde{f}(P_1, P_2, \ldots, P_n) = f(P_1, P_2, \ldots, P_n, \bar{P}_{n+1}).$$

Since f is strategy-proof, it is clear that also $\tilde{f}$ must be strategy-proof. Moreover, since $\mathscr{O}_{-(n+1)}(\bar{P}_{n+1}, f) \cap \{a_1, a_2, a_3\} = \{a_1, a_2, a_3\}$, it follows from Observation 1 that $\tilde{f}$ satisfies unanimity with respect to a_1, a_2 and a_3. Thus, we can apply the induction hypothesis to $\tilde{f}$ and conclude that there exists some individual i such that

$$\mathscr{O}_{-i}(P_i, \tilde{f}) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\} \qquad \text{for all} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3. \tag{4.11}$$

Consider now a preference profile $\mathscr{P} = (P_1, P_2, \ldots, P_n)$ such that $r_1(P_i) = a_2$ and $r_1(P_j) = a_1$ for all $j = 1, 2, \ldots, n$ and $j \neq i$. By (4.11), we have $a_1 \notin \mathscr{O}_{-i}(P_i, \tilde{f})$, and hence $\tilde{f}(P_1, P_2, \ldots, P_n) \neq a_1$. But this means that $f(P_1, P_2, \ldots, P_n, \bar{P}_{n+1}) \neq a_1$, and we conclude that $a_1 \notin \mathscr{O}_{-i}(P_i, f)$ because otherwise we would obtain a contradiction to Observation 1. By the previous lemma follows now that

$$\mathscr{O}_{-i}(P_i, f) \cap \{a_1, a_2, a_3\} = \{r_1(P_i)\} \qquad \text{for all} \quad P_i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3,$$

which completes the induction step, and the lemma is proved □

From now on we will, without loss of generality, assume that the unique individual satisfying (4.8) is individual 1. The previous lemma makes a negative statement by stating that if $r_1(P_1) = a_1$, then $f(P_1, P_{-1}) \neq a_2, a_3$, but it does not tell us what the social choice actually will be, since $f(P_1, P_{-1})$ still could belong to $\mathscr{A} \setminus \{a_1, a_2, a_2\}$. The next lemma shows, however, that the social choice in deed will be individual 1's top alternative, at least when there is complete unanimity in the rest of the society. To simplify the argumentation, we introduce the function $f_1 : \Gamma^2 \to \mathscr{A}$, defined by

$$f_1(P_1, P_2) = f(P_1, P_2, \ldots, P_2), \tag{4.12}$$

which means that f_1 reports the social choice of f when individual 1 chooses P_1 and the rest of the society agrees unanimously on P_2. Note that f_1 is top-2 strategy-proof according to Lemma 4.4 because f is top-2 strategy-proof.

Lemma 4.9. *Suppose that*

$$\mathscr{O}_{-1}(P_1) \cap \{a_1, a_2, a_3\} = \{r_1(P_1)\} \qquad \textit{for all} \quad P_1 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3.$$

Then for all $P_1, P_2 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$*, we have*

$$f_1(P_1, P_2) = r_1(P_1).$$

Proof. When $r_1(P_1) = r_1(P_2)$, the claim follows from the assumption that f satisfies unanimity with respect to a_1, a_2 and a_3. Suppose therefore, without loss of generality, that $r_1(P_1) = a_1$ and $r_1(P_2) = a_2$. If $r_2(P_1) = a_2$, Observation 2 implies $f_1(P_1, P_2) = a_1$, and from the monotonicity property of top-2 strategy-proof social choice functions follows then that $f(P_1, P_2) = a_1$ for all $P_1 \in \Gamma_1$. □

We turn now to the second part of the proof, which is mainly an induction step where we show that if individual 1's choice P_1 and the unanimous choice P_2 of the rest of the society have their top alternatives among $\{a_1, a_2, \ldots, a_k\}$ and f chooses individual 1's top alternative, then the same holds when the set of possible top alternatives is extended to $\{a_1, a_2, \ldots, a_k, a_{k+1}\}$. This induction step is carried out in the next two lemmas, and in the first one we will only allow P_2 to extend its set of possible top alternatives. From now on we will assume that f satisfies unanimity with respect to all $a \in \mathscr{A}$.

Lemma 4.10. *Suppose that for all* $P_1, P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k$ *we have*

$$f_1(P_1, P_2) = r_1(P_1). \tag{4.13}$$

Then (4.13) *also holds for all* $P_1 \in \Gamma_1 \cup \ldots \cup \Gamma_k$ *and* $P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k \cup \Gamma_{k+1}$.

Proof. By assumption, (4.13) holds already for $P_1, P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k$, so it suffices to consider the case when $P_2 \in \Gamma_{k+1}$ and $P_1 \in \Gamma_1 \cup \ldots \cup \Gamma_k$. For P_1, we will consider one of the sets Γ_i for $1 \leq i \leq k$ at a time, and we will prove the lemma using induction over the number of connections between a_i and a_{k+1}. Suppose therefore first that $a_i \sim a_{k+1}$. Since Γ is linked there exists some $a_j \in \{a_1, a_2, \ldots, a_k\}$ such that $a_j \sim a_{k+1}$ and $a_j \neq a_i$. Consider first the preferences

P_1	P_2	P_2'
a_i	a_{k+1}	a_j
a_{k+1}	$\vdots$	a_{k+1}

By simple Pareto optimality, we have $f_1(P_1, P_2) \in \{a_i, a_{k+1}\}$. On the other hand, since $P_2' \in \Gamma_j$, we get from (4.13) that $f_1(P_1, P_2') = a_i$. Hence, if $f_1(P_1, P_2) = a_{k+1}$, then individual 2 would be able to manipulate f_1 by going from (P_1, P_2') to (P_1, P_2). Thus, $f_1(P_1, P_2) = a_i$, and by monotonicity we conclude then that $f_1(P_1, P_2) = a_i$ for all $P_1 \in \Gamma_i$.

Suppose next that the lemma already has been proved for n connections, and let $[a_i, a_{i_1}, a_{i_2}, \ldots, a_{i_n}, a_{k+1}]$ be a chain of $n+1$ connections, where $a_{i_1}, a_{i_2}, \ldots, a_{i_n} \in \{a_1, a_2, \ldots, a_k\}$. Consider first the preferences

P_1	P_1'	P_2	P_2'
a_i	a_{i_1}	a_{k+1}	a_{i_1}
a_{i_1}	a_i	$\vdots$	a_i

By the induction hypothesis, we have $f_1(P_1', P_2) = a_{i_1}$. Since f_1 is top-2 strategy-proof we can therefore conclude that $f_1(P_1, P_2) \in \{a_i, a_{i_1}\}$. On the other hand, by (4.13) we have $f_1(P_1, P_2') = a_i$. Thus, if $f_1(P_1, P_2) = a_{i_1}$, then individual 2 would be able to manipulate f_1 by moving from (P_1, P_2') to (P_1, P_2). Hence, $f_1(P_1, P_2) = a_i$, and by monotonicity follows then that $f_1(P_1, P_2) = a_i$ for all $P_1 \in \Gamma_i$, and the lemma is proved. □

In the following second lemma of the induction step, also individual 1 is allowed to pick his top alternative from the extended set $\{a_1, a_2, \ldots, a_k, a_{k+1}\}$.

Lemma 4.11. *Suppose that for all $P_1 \in \Gamma_1 \cup \ldots \cup \Gamma_k$ and $P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k \cup \Gamma_{k+1}$ we have*

$$f_1(P_1, P_2) = r_1(P_1). \tag{4.14}$$

Then (4.14) *holds also for all $P_1, P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k \cup \Gamma_{k+1}$.*

Proof. By assumption (4.14) holds already for $P_1, P_2 \in \Gamma_1 \cup \ldots \cup \Gamma_k$, whence it suffices to consider the case when $P_1 \in \Gamma_{k+1}$ and $P_2 \in \Gamma_i$ for some i with $1 \leq i \leq k+1$. When $i = k+1$, then $f_1(P_1, P_2) = a_{k+1}$ by unanimity, and there is nothing to prove. Suppose therefore that $1 \leq i \leq k$. The lemma will now be proved by separately considering the two cases when a_i and a_{k+1} are connected, respectively when they are not connected. Suppose therefore first that $a_i \sim a_{k+1}$. Let $a_j \in \{a_1, a_2, \dots, a_k\}$ be such that $a_j \sim a_{k+1}$ and $a_j \neq a_i$, and consider the preferences

P_1	P_1'	P_1''	P_1'''	P_2	P_2'
a_{k+1}	a_{k+1}	a_{k+1}	a_j	a_i	a_i
$\vdots$	a_i	a_j	a_{k+1}	$\vdots$	a_{k+1}

We will use a contradiction argument and assume that $f_1(P_1, P_2) \neq a_{k+1}$. Then, by strategy-proofness, we also must have $f_1(P_1', P_2) \neq a_{k+1}$. But by simple Pareto optimality $f_1(P_1', P_2) \in \{a_i, a_{k+1}\}$, and hence $f_1(P_1', P_2) = a_i$. Applying monotonicity to (P_1', P_2), we conclude that then also $f_1(P_1', P_2') = a_i$. But then we must also have $f_1(P_1'', P_2') = a_i$; this holds since $f_1(P_1'', P_2') \in \{a_i, a_{k+1}\}$ by simple Pareto optimality, but strategy-proofness excludes a_{k+1}, because otherwise individual 1 would be able to top-2 manipulate f_1 at (P_1', P_2'). However, by (4.14) we have $f_1(P_1''', P_2') = a_j$, so individual 1 can top-2 manipulate from (P_1'', P_2') to (P_1''', P_2'), which contradicts strategy-proofness. Thus, our initial assumption must have been wrong, and we conclude that $f_1(P_1, P_2) = a_{k+1}$ for all $(P_1, P_2) \in \Gamma_{k+1} \times \Gamma_i$ with $a_{k+1} \sim a_i$.

We turn now to the case when a_i and a_{k+1} are not connected. Choose again $a_j \in \{a_1, a_2, \dots, a_k\}$ such that $a_j \sim a_{k+1}$ and $a_j \neq a_i$, and consider the preferences

P_1	P_1'	P_1''	P_2	P_2'
a_{k+1}	a_{k+1}	a_j	a_i	a_j
$\vdots$	a_j	a_{k+1}	$\vdots$	$\vdots$

By (4.14) we have $f_1(P_1'', P_2) = a_j$. Since f_1 is top-2 strategy-proof, we can conclude that $f_1(P_1', P_2) \in \{a_j, a_{k+1}\}$, because otherwise individual 1 would be able to obtain his second best alternative by going from (P_1', P_2) to (P_1'', P_2). By the previous paragraph, we have $f_1(P_1', P_2') = a_{k+1}$, since $a_j \sim a_{k+1}$. But then we must also have $f_1(P_1', P_2) = a_{k+1}$, because if $f_1(P_1', P_2) = a_j$, then individual 2 would be able to manipulate f_1 by going from (P_1', P_2') to (P_1', P_2). But $f_1(P_1', P_2) = a_{k+1}$ implies by monotonicity that $f_1(P_1, P_2) = a_{k+1}$, and the lemma is proved. □

We can now summarize the preceding chain of lemmas and prove Theorem 4.1.

Proof of Theorem 4.1. Let $f : \Gamma^N \to \mathscr{A}$ be a social choice function that satisfies unanimity. If f is dictatorial, then f is obviously top-2 strategy-proof. Conversely, suppose that f is top-2 strategy-proof and we want to prove that f then must be dictatorial. Since f satisfies unanimity, f satisfies in particular unanimity with respect to a_1, a_2 and a_3. By Lemma 4.8, we can therefore conclude that there exists exactly one $i \in \mathcal{I}$ such that (4.8) holds, and for simplicity we assume that $i = 1$. According to Lemma 4.9, the function $f_1 : \Gamma^2 \to \mathscr{A}$ defined in (4.12) satisfies

$$f_1(P_1, P_2) = r_1(P_1) \qquad \text{for all} \quad P_1, P_2 \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3.$$

But then we can use Lemma 4.10 and Lemma 4.11 repeated times to conclude that $f_1(P_1, P_2) = r_1(P_1)$ still holds when we successively extend the domain of P_1 and P_2. Noting that $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_M$, we finally get

$$f_1(P_1, P_2) = r_1(P_1) \qquad \text{for all} \quad P_1, P_2 \in \Gamma. \tag{4.15}$$

It remains to show that (4.15) implies that individual 1 is a dictator for f. We will argue by contradiction and suppose therefore that for some preference profile $(P_1, P_2, \ldots, P_N) \in \Gamma^N$ we have

$$f(P_1, P_2, \ldots, P_N) = a \qquad \text{and} \qquad a \neq r_1(P_1).$$

Since Γ is linked, there must be a preference $P_2' \in \Gamma$ with $r_1(P_2') = a$ because a is connected with at least two alternatives. Replacing all preferences except from the first one in $(P_1, P_2, \ldots, P_N)$ by P_2' we obtain by monotonicity that

$$f_1(P_1, P_2') = f(P_1, P_2', \ldots, P_2') = a \neq r_1(P_1).$$

As this contradicts (4.15), we conclude that individual 1 must be a dictator for f, and the proof of Theorem 4.1 is fulfilled. □

We finish this subsection with some remarks on the history of the preceding proof. The idea to prove the impossibility of strategy-proof social choice by an analysis of option sets has its origin in Barberà and Peleg (1990). They introduced the notion of option sets and used it to prove the Gibbard-Satterthwaite theorem for the case when the number of alternatives in $\mathscr{A}$ is infinite. After that, a similar technique was used in Aswal et al. (2003) to prove that a strategy-proof social choice function on a linked domain must be dictatorial (see Theorem 3.2 in this thesis). Theorem 4.1 is then a natural generalization of this result to linked top-2 domains, and of course, our proof of Theorem 4.1 is to a great extent inspired by the corresponding proof in Aswal et al. (2003), whence we will give some details about how these two proofs are related. Aswal et al. (2003) use an induction argument in their proof: first, they prove their theorem for $N = 2$, and second, they show that if their theorem holds for $N = 2$, then it actually holds for all $N \geq 2$. Concerning the first step, I am convinced that the proof in Aswal et al. (2003) can, with minor modifications, also be applied to linked top-2 domains. The induction step, however, cannot be applied to top-2 domains, and since it may be instructive to see why, we present it in Appendix B and explain why it fails for partial preference relations. Because we were not able to find an alternative proof for the induction step for linked top-2 domains, we endeavoured to generalize the proof in Aswal et al. (2003) for $N = 2$ to a general N. In deed, Lemma 4.5 to Lemma 4.7 in the proof above are more or less the same as Lemma 3.1 to Lemma 3.3 in Aswal et al. (2003). Thereafter, however, our proof deviates from the proof in Aswal et al. (2003) in a fundamental way.

4.3 The Gibbard-Satterthwaite Theorem and Top-2 Manipulability

We will now as a first application of Theorem 4.1 state a more informative version of the Gibbard-Satterthwaite theorem. If $\mathscr{A}$ is a finite set of alternatives, then the set Σ of all complete, antisymmetric and transitive preferences over $\mathscr{A}$ is obviously a linked top-2 domain, and applying Theorem 4.1 to Σ, we get:

Theorem 4.12 (Strengthening of the Gibbard-Satterthwaite Theorem). *If $\mathscr{A}$ is a finite set of at least three alternatives, then every non-dictatorial social choice function $f : \Sigma^N \to \mathscr{A}$ that satisfies unanimity is top-2 manipulable.*

Theorem 4.12 tells us that if $f : \Sigma^N \to \mathscr{A}$ is a non-dictatorial social choice function that satisfies unanimity, then some individual can at some preference profile manipulate f and obtain by this at least his second best alternative. It is now natural to ask whether this result can be strengthened further in the sense that every social choice function $f : \Sigma^N \to \mathscr{A}$ that satisfies unanimity is not only top-2 manipulable, but that some individual by manipulation at some preference profile actually can obtain his top alternative. The answer to this question, however, is negative in general because for example the majority rule is such that no voter ever can obtain his top alternative be manipulation.[18] By this, Question 4 from the introduction is answered.

[18]This can be seen in the following way: If an individual's top alternative gets most votes when the individual misrepresents his preferences and does not vote for this alternative, then this alternative gets surely also most votes when the individual votes for his top alternative. Thus, if misrepresentation gives an individual his top alternative, then also sincere voting does so, and therefore, a top alternative can never be obtained by manipulation.

5 Strategy-proof Social Choice of Fixed-sized Subsets

WE ARE NOW sufficiently prepared to investigate to what extent strategy-proof social choice of fixed-sizes subsets is possible by answering the first three questions proposed in the introduction. To begin with, we formalize in Section 5.1 the social choice of fixed-sized subsets and derive a preference structure that is reasonable in this context. After this, we show in Section 5.2 that non-dictatorial strategy-proof social choice of fixed-sized subsets in analogy with the classical Gibbard-Satterthwaite theorem is impossible in general. This result is then modified in Section 5.3, where we show that the social choice of fixed-sized subsets can be made strategy-proof when voters' preferences are single-peaked. In Section 5.4, finally, we compare the voting model and the results presented in this chapter with related contributions in the literature of strategy-proof social choice theory.

5.1 Strategy-proof Social Choice of Subsets Based on Preferences over Alternatives

In this thesis, the social choice of fixed-sized subsets will be considered in the following formal framework: A society consisting of N individuals has to choose k elements from a set $\mathscr{A}$ that contains M alternatives. The set of all possible social choices, i.e., the set of all subsets of $\mathscr{A}$ that contain exactly k elements, will be denoted by $\mathscr{A}_k$. We will assume that the individuals in the society have complete, antisymmetric and transitive preferences over the alternatives in $\mathscr{A}$, and the set of all such preferences will be denoted by Σ, as before. We assume further that social choice is based on these preferences by using a social choice function of the form

$$f : \Sigma^N \to \mathscr{A}_k. \tag{5.1}$$

Note that even though society is going to choose an element from $\mathscr{A}_k$, we assume that the arguments of the social choice function f are the individuals' preferences over $\mathscr{A}$, but not over $\mathscr{A}_k$, which is motivated by the fact that in most of the practically used voting procedures for the social choice of fixed-sized subsets, voters are only required to report (a part of) their preferences over the alternatives in $\mathscr{A}$ on the ballots, but not over subsets of $\mathscr{A}$. We illustrate by the following two examples how a social choice function for the social choice of fixed-sized subsets can look like in practice, and we will return to these examples later in this chapter.

Example 5.1. A social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$ can easily be obtained by generalizing the majority rule from Example 2.1: Let each voter report his most preferred alternative in $\mathscr{A}$, and let the social choice be the subset of $\mathscr{A}$ that contains the k alternatives that got most votes. Similarly, we can generalize the Borda count from Example 2.2: Let each voter assign points to the alternatives in $\mathscr{A}$, and let the social choice be the subset of $\mathscr{A}$ that contains the k alternatives that got most points in total. □

Example 5.2. A rather different procedure is used to elect the 659 members in the British parliament, the House of Commons: The United Kingdom is divided into 659 electoral districts, and in each of them, voters can choose among a number of candidates. The candidate who receives most votes in a district will take a seat in the parliament. □

We noted already in the introduction that it is not obvious when a social choice function of the form (5.1) should be considered manipulable, because contrary to the case of the classical Gibbard-Satterthwaite theorem, where the social choice function is of the form $f : \Sigma^N \to \mathscr{A}$, the preferences in Σ cannot be used directly to rank the elements in $\mathscr{A}_k$, whence it is not clear when an individual that is able to change the social choice by misrepresentation gains from doing so. In the following, we will therefore answer Question 1 from the introduction and show how the preferences in Σ can be transferred to preferences over $\mathscr{A}_k$, which then automatically leads to a notion of manipulability for a social choice function of the form (5.1). To begin with, we illustrate why an individual can find it profitable to misrepresent his preference when the generalized ordinary majority rule from Example 5.1 is used.

Example 5.3. Using the generalized ordinary majority rule from Example 5.1, there might be two reasons for a voter to misrepresent his preferences: On the one hand, when the voter realizes that his top alternative has too little support among other voters to be elected, then he might want to give his vote to an alternative that has a more reasonable chance to be elected. On the other hand, when the voter can be sure that his top alternative has so broad support among other voters that it surely will be elected, then he might think about voting for another alternative that he also would like to see elected. □

The two possible motives for misrepresentation in Example 5.3 are based on the same underlying reasoning: By casting his vote on an alternative other than his most preferred one, a voter hopes to lift in a desirable alternative into the socially chosen subset, thereby pushing out a less desirable alternative. In general, we will therefore assume that individual preferences satisfy the following *separability condition*: If A_1 and A_2 are two subsets of $\mathscr{A}$ containing exactly k elements, then an individual will prefer A_1 to A_2 if A_1 can be obtained from A_2 by replacing one of the alternatives in A_2 by an alternative that the individual finds more desirable. Formally, if $\bar{A}$ is a subset of $\mathscr{A}$ containing $k-1$ elements, and $A_1 = \bar{A} \cup \{a\}$ and $A_2 = \bar{A} \cup \{a'\}$, then

$$A_1 \text{ is preferred to } A_2 \qquad \Longleftrightarrow \qquad a \text{ is preferred to } a'. \tag{5.2}$$

In addition to (5.2), we will of course also assume that an individual prefers a subset $A_1 \in \mathscr{A}_k$ to a subset $A_2 \in \mathscr{A}_k$ if A_1 can be obtained from A_2 by successive replacement of several of the elements in A_2 by better alternatives. On the other hand, if it is not possible to obtain one of A_1 or A_2 from the other one in this manner, then the only way to transfer A_1 to A_2 is to replace some of the elements in A_1 by better and some by worse alternatives, so with available information it is not possible to decide unambiguously which subset an individual will prefer, and we will therefore in this case assume that an individual has no explicit preference between A_1 and A_2. In the described way, we can associate to every $P \in \Sigma$ a unique preference $\hat{P}$ on $\mathscr{A}_k$, which we will call the *preference induced by P on $\mathscr{A}_k$*, and the formal definition of this preference is presented below.

Definition 5.1 (Induced Preference on $\mathscr{A}_k$). For $P \in \Sigma$, the *preference $\hat{P}$ induced by P on $\mathscr{A}_k$* is defined as follows: If $A_1, A_2 \in \mathscr{A}_k$, then $A_1 \hat{P} A_2$ if and only if $A_1 \neq A_2$, and we can decompose A_1 and A_2 as

$$A_1 = \bar{A} \cup \{a_{i_1}\} \cup \{a_{i_2}\} \cup \ldots \cup \{a_{i_l}\}$$
$$A_2 = \bar{A} \cup \{a_{i'_1}\} \cup \{a_{i'_2}\} \cup \ldots \cup \{a_{i'_l}\}$$

where $\bar{A} = A_1 \cap A_2$, and we have $a_{i_1} P a_{i'_1}$, $a_{i_2} P a_{i'_2}$, $\ldots$, and $a_{i_l} P a_{i'_l}$. The set of all induced preferences over $\mathscr{A}_k$ will be denoted by $\Gamma(\mathscr{A}_k)$.

Definition 5.1 is our answer to Question 1 from the introduction. Note that if $\hat{P}$ is induced on $\mathscr{A}_k$ by some $P \in \Sigma$, then $\hat{P}$ is *antisymmetric* and *transitive*,[19] and we can thus apply the notions of partial preferences introduced in Chapter 4 to $\hat{P}$. We

[19] Antisymmetry follows directly from the definition of $\hat{P}$, and transitivity can be proved as follows: Suppose that $A_1, A_2, A_3 \in \mathscr{A}_k$ are such that $A_1 \hat{P} A_2$ and $A_2 \hat{P} A_3$. Then A_2 can be obtained from A_3 by replacing some of the elements in A_3 by alternatives that are better according to P, and similarly, replacing some of the elements in A_2 by better alternatives, we can get from A_2 to A_1. But this means of course that if all replacements are carried out at one sweep, then we can get from A_3 to A_1 by replacing worse alternatives by better alternatives, and hence $A_1 \hat{P} A_3$.

illustrate by the following example how Definition 5.1 works in a concrete situation, and this example shows also that $\hat{P}$ is in general *not complete*.[20]

Example 5.4. Let $\mathscr{A} = \{a, b, c, d, e\}$ be a set of alternatives, and suppose that the preference P ranks the elements in $\mathscr{A}$ according to $a\,P\,b\,P\,c\,P\,d\,P\,e$, and let $\hat{P}$ be the preference induced by P on $\mathscr{A}_3$ according to Definition 5.1. Assume that society has to elect three of the alternatives in $\mathscr{A}$, and consider the four subsets

$$A_1 = \{a,b,c\}, \quad A_2 = \{a,c,e\}, \quad A_3 = \{c,d,e\}, \quad \text{and} \quad A_4 = \{b,c,d\},$$

which, of course, all belong to $\mathscr{A}_3$. We note that A_1 can be obtained from A_2 by replacing alternative e by b, and since $b\,P\,e$, we have $A_1\,\hat{P}\,A_2$. Similarly, replacing d in A_3 by a, we obtain A_2, and thus $A_2\,\hat{P}\,A_3$, because $a\,P\,d$. We also have $A_1\,\hat{P}\,A_3$, because A_1 can be obtained from A_3 by replacing d by a and e by b, which indicates that $\hat{P}$ is transitive.

But consider now A_2 and A_4. On the one hand, A_2 contains a, which is preferred to all elements in A_4, but on the other hand, A_2 also contains e, to which all elements in A_4 are preferred. Hence, A_2 and A_4 are not comparable by $\hat{P}$, which shows that the preference $\hat{P}$ induced by P on $\mathscr{A}_k$ will not be complete in general. □

Remark 5.1. It must be pointed out that even if the equivalence in (5.2), on which Definition 5.1 is based, seems very reasonable, it is nevertheless an *assumption*. Verbally, (5.2) means that a voter's opinion on which of two alternatives should be included in a subset is independent of which other alternatives are contained in the subset. This separability assumption, however, can fail for at least two reasons.

Firstly, there might be dynamic effects between the alternatives, which can affect an individual's ranking of different subsets of $\mathscr{A}$. Suppose, for example, that you as a head of an institute have to appoint a research group consisting of five members, of which four already are elected, and to fill the remaining place, you can choose between doctor Jekyll and professor Hyde. If you think that doctor Jekyll is a much more skilled researcher than professor Hyde, but professor Hyde is able to co-operate much more efficiently with the other members in the research group, and you will therefore appoint professor Hyde, then your preferences do not satisfy (5.2).

Secondly, an individual might form his opinion of a subset of $\mathscr{A}$ not only depending on which alternatives are included in the subset, but also on the structure of the subset as a whole. This is the case, for instance, when the individual considers

[20] There are actually three instances when $\hat{P}$ is complete, namely when k equals 1, $M-1$, or M.

that a balanced distribution of sex or age among the members in a committee is important, and then the individual's preferences over subsets of $\mathscr{A}$ will in general not satisfy (5.2). □

Via Definition 5.1, preferences over $\mathscr{A}$ can be transferred to preferences over $\mathscr{A}_k$, and it is now straight forward to define a notion of manipulability for a social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$.

Definition 5.2 (Manipulation of the Social Choice of Fixed-sized Subsets). A social choice function $f : \Sigma^N \to \mathscr{A}_k$ is said to be *manipulable* if there exist $P_i, P_i' \in \Sigma$ and $P_{-i} \in \Sigma^{N-1}$ such that

$$f(P_i', P_{-i}) \ \hat{P}_i \ f(P_i, P_{-i}),$$

where $\hat{P}_i$ is the preference induced by P_i on $\mathscr{A}_k$.

Having a notion of manipulability, it makes now sense to ask whether there exist any strategy-proof social choice functions of the form $f : \Sigma^N \to \mathscr{A}_k$, and we will turn to this question in Section 5.2 and Section 5.3.

Before leaving this section, we would like to remark that Definition 5.1 is not the only way to define preferences over $\mathscr{A}_k$ based on preferences over $\mathscr{A}$, and we will therefore briefly consider how one can proceed alternatively. In order to define a preference $\bar{P}$ on $\mathscr{A}_k$ based on a preference P on $\mathscr{A}$, we think that it is almost imperative to require that if $A_1, A_2 \in \mathscr{A}_k$ are such that A_1 can be obtained from A_2 by replacing some of the elements in A_2 by more desirable alternatives, then we must have $A_1 \bar{P} A_2$, or with other words, $\bar{P}$ should in any case be compatible with $\hat{P}$. If none of A_1 and A_2 can be obtained from the other by replacing worse alternatives by better alternatives, then we abstained in our approach from regarding one of A_1 or A_2 to be preferred to the other, because we have no additional information that allows us to do so. This approach, however, is not unproblematic, because the preference $\hat{P}$ defined in this way is in general not complete, which makes the mathematical analysis of whether a social choice function $f : \Sigma^N \to \mathscr{A}_k$ can be strategy-proof more complicated, and since there to our knowledge exist no results for strategy-proof social choice functions with incomplete preferences in the literature, we were forced to derive the notions and results in Chapter 4 in order to fulfil the purpose of this thesis. To avoid the problems caused by the incompleteness of $\hat{P}$, one could thus have thought of extending $\hat{P}$ to a complete preference $\bar{P}$ over $\mathscr{A}_k$, which according to Szpilrajn's theorem always is possible, and we present below two thinkable ways to do so. Note, however, that the approach chosen in this thesis is the more general one, because we can show that a non-dictatorial social choice function $f : \Sigma^N \to \mathscr{A}_k$ must be manipulable with respect to the preferences defined in Definition 5.1, and therefore f must of course also be manipulable with respect to preferences that are compatible with the preferences from Definition 5.1, but the converse need not be true.

Firstly, the possibly most natural approach to solve the problems caused by the incompleteness of $\hat{P}$ is to interpret the absence of strict preference between subsets that are not ranked by $\hat{P}$ as *indifference*. To be precise, based on $\hat{P}$, we would like to define a complete, but weak preference $\hat{R}$ on $\mathscr{A}_k$ by

i) if $A_1, A_2 \in \mathscr{A}_k$ and $A_1 \hat{P} A_2$, then $A_1 \hat{R} A_2$, and

ii) if $A_1 = A_2$ or A_1 and A_2 are not ranked by $\hat{P}$, then both $A_1 \hat{R} A_2$ and $A_2 \hat{R} A_1$.

The preference $\hat{R}$ defined in this way is certainly complete and compatible with the preference $\hat{P}$. However, $\hat{R}$ has also a severe weakness because it is not transitive in general. To see this, set $\mathscr{A} = \{a_1, a_2, a_3, a_4\}$, and suppose that the preference P orders the alternatives in $\mathscr{A}$ according to $a_1 P a_2 P a_3 P a_4 P a_5$. Let further $\hat{P}$ be the preference induced by P on $\mathscr{A}_k$, and suppose that $\hat{R}$ is defined as above. According to the definition of $\hat{R}$, the following two implications must hold:

$$\begin{aligned} &\{a_1, a_5\} \text{ and } \{a_2, a_3\} \text{ are not ranked by } \hat{P} \quad \Longrightarrow \quad \{a_1, a_5\} \, \hat{R} \, \{a_2, a_3\}, \\ &\{a_2, a_3\} \text{ and } \{a_1, a_4\} \text{ are not ranked by } \hat{P} \quad \Longrightarrow \quad \{a_2, a_3\} \, \hat{R} \, \{a_1, a_4\}. \end{aligned} \tag{5.3}$$

If $\hat{R}$ would have been transitive, then (5.3) would imply $\{a_1, a_5\} \hat{R} \{a_1, a_4\}$, but this cannot be the case because $\{a_1, a_4\} \hat{P} \{a_1, a_5\}$. This lack of transitivity indicates that the absence of strict preference is not equivalent to indifference, and therefore we reject this approach.

Secondly, a more successful approach is to assume that the individuals in the society base their preferences over $\mathscr{A}_k$ on *cardinal preferences* over $\mathscr{A}$. By this, we mean that the individuals assign to each alternative in $\mathscr{A}$ a real number, a *utility*, in such a way that the more desirable an alternative is, the higher is its utility, and an individual prefers a subset $A_1 \in \mathscr{A}_k$ to $A_2 \in \mathscr{A}_k$ if and only if the sum of the utilities of the elements in A_1 is larger than the corresponding sum for A_2. Interpretating equality of two sums as indifference between the corresponding subsets, we obtain preferences over $\mathscr{A}_k$ that are complete, transitive, and compatible with $\hat{P}$ because replacing an element in a subset with a more desirable alternative increases of course the sum of utilities. Furthermore, with almost the same arguments which will be applied to $\hat{P}$ in the next section, it is possible to show that the set of all preferences over $\mathscr{A}_k$ that are constructed in this way is a linked domain, and we could therefore apply Theorem 3.2 to $f : \Sigma^N \to \mathscr{A}_k$.[21] However, one can object against this approach that it is more unrealistic than that chosen in this thesis, because real people do not likely form their opinions of subsets of $\mathscr{A}$ by summing up utilities. On the contrary, Definition 5.1 seems to be a good approximation of the mental process by which people actually compare subsets of $\mathscr{A}$.

5.2 The Case of Unrestricted Preferences

After having derived a notion of manipulability for a social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$ in the previous section, we turn now to Question 2 from the introduction and investigate whether there exist any non-dictatorial strategy-proof social choice functions for the social choice of fixed-sized subsets when the domain of preferences is assumed to be unrestricted. Like in the case of the original Gibbard-Satterthwaite theorem, it turns out that strategy-proofness and non-dictatorship do not exclude each other in general, but under an additional efficiency requirement, strategy-proofness implies dictatorship. To begin with, we show by an example how one can construct a non-dictatorial strategy-proof social choice function for the social choice of fixed-sized subsets when some degree of inefficiency is accepted.

Example 5.5. Consider a simplified variant of the voting procedure used in the elections to the House of Commons presented in Example 5.2: Suppose that society has to choose two of the alternatives in the set $\mathscr{A} = \{a_1, a_2, a_3, a_4\}$, and to achieve this, voters are divided into two groups, of which the first one chooses one of the elements in $\{a_1, a_2\}$, and the second one chooses one of the elements in $\{a_3, a_4\}$. If voters' preferences satisfy (5.2), then this voting procedure is obviously strategy-proof because a voter belonging to the first group, for example, must take the second group's choice as given, and the best thing he can do when choosing between a_1 and a_2 is therefore to vote sincerely. However, this procedure has an unappealing lack of efficiency. Suppose, for instance, that all voters unanimously agree on that both a_1 and a_2 are better than both a_3 and a_4. Then society's natural choice should be $\{a_1, a_2\}$, but this can never be the outcome of this procedure. □

Consequently, strategy-proofness without an efficiency requirement does not imply

[21] Actually, even with this approach, Theorem 3.2 must be adjusted because preferences over $\mathscr{A}_k$ that are based on cardinal preferences are not necessarily strict. However, they turn out to have a unique first and a unique second alternative in $\mathscr{A}_k$, and an investigation of the proof of Theorem 3.2 in Aswal et al. (2003) shows that this theorem actually also holds for weak preferences of this type.

dictatorship. Recall that in the original Gibbard-Satterthwaite theorem, the social choice function $f : \Sigma^N \to \mathscr{A}$ is assumed to be efficient in the sense that it satisfies *unanimity*, which simply means that if $r_1(P_i) = a$ for all $i \in \{1,2,\ldots,N\}$, then $f(P_1,P_2,\ldots,P_N) = a$. It seems natural to assume a corresponding condition for a social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$, but this is not entirely straightforward because there are two plausible ways to generalize the notion of unanimity: Supposed that society has to choose k of the alternatives in $\mathscr{A}$, we will say that there is *complete unanimity* in the society when all individuals agree exactly on the ranking of the k best alternatives, that is, all individuals have the same top alternative, the same second best alternative, and so on, up to the kth ranked alternative,[22] and we will say that there is *weak unanimity* in the society whenever all individuals agree on that k of the alternatives in $\mathscr{A}$ are preferred to all other alternatives, but they may disagree on the internal order of these k alternatives.[23] A social choice function $f : \Sigma^N \to \mathscr{A}_k$ is of course said to satisfy complete respectively weak unanimity if the value of f equals that subset of $\mathscr{A}$ that contains the k alternatives that are unanimously preferred to all other alternatives whenever there is complete respectively weak unanimity in the society. Complete unanimity is obviously a special case of weak unanimity, and therefore, a social choice function that satisfies weak unanimity satisfies also complete unanimity, but the converse need not be true in general. However, it turns out that if the social choice function is strategy-proof, then also the converse implication holds.[24] Hence, it makes no difference which notion of unanimity we use, and we will therefore use the apparently weaker assumption that the social choice function satisfies complete unanimity. Under this condition, it turns out that strategy-proofness implies dictatorship:

Theorem 5.1 (The Gibbard-Satterthwaite theorem for the social choice of fixed-sized subsets). *Suppose that $\mathscr{A}$ is a set of $M \geq 3$ alternatives, and let $f : \Sigma^N \to \mathscr{A}_k$, where $1 \leq k \leq M-1$, be a social choice function that satisfies complete unanimity. Then f is strategy-proof if and only if f is dictatorial.*

[22] Formally, there is *complete unanimity* in the society if $r_1(P_i) = a_1$, $r_2(P_i) = a_2, \ldots, r_k(P_i) = a_k$ for all $i \in \{1,2,\ldots,N\}$ and some $a_1, a_2, \ldots, a_k \in \mathscr{A}$.

[23] Formally, there is *weak unanimity* in the society if there exists a subset $A \in \mathscr{A}_k$ such that $\{r_1(P_i), r_2(P_i), \ldots, r_k(P_i)\} = A$ for all $i \in \{1,2,\ldots,N\}$.

[24] To see this, suppose that $f : \Sigma^N \to \mathscr{A}_k$ is strategy-proof, and for all $i \in \{1,2,\ldots,N\}$ we have $\{r_1(P_i), r_2(P_i), \ldots, r_k(P_i)\} = A$ for some $A \in \mathscr{A}_k$, but $f(P_1,P_2,\ldots,P_N) \neq A$. However, we must have $f(P_1,P_1,\ldots,P_1) = A$, which means that if we first replace P_2 by P_1, next P_3 by P_1, and so on up to P_N, then there must be some individual that changes the value of f to A, and hence manipulates f. Thus, a strategy-proof social choice function that satisfies complete unanimity satisfies also weak unanimity.

The rest of this section is devoted to a formal proof of Theorem 5.1. As mentioned before, Theorem 5.1 will be proved using Theorem 4.1 from Chapter 4, but it is not possible to apply Theorem 4.1 directly because the social choice function in Theorem 4.1 is such that its arguments are preferences over the set of all possible social outcomes, and a social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$ does not have this property. Therefore, we need to adjust the arguments in f to be preferences over $\mathscr{A}_k$. Note that the preferences in Σ and $\Gamma(\mathscr{A}_k)$ are in a one-to-one correspondence in the sense that not only every $P \in \Sigma$ induces a unique $\hat{P} \in \Gamma(\mathscr{A}_k)$, but that also for every $\hat{P} \in \Gamma(\mathscr{A}_k)$ exists a unique $P \in \Sigma$ that induces $\hat{P}$, because P can via (5.2) be entirely re-constructed from $\hat{P}$. To every social choice function $f : \Sigma^N \to \mathscr{A}_k$, we can thus uniquely associate a social choice function $\hat{f} : \Gamma(\mathscr{A}_k)^N \to \mathscr{A}_k$ by

$$\hat{f}(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_N) \stackrel{\text{def}}{=} f(P_1, P_2, \ldots, P_N), \tag{5.4}$$

where P_i for all $i \in \{1, 2, \ldots, N\}$ is that preference in Σ that induces $\hat{P}_i$. It is obvious that $\hat{f}$ is dictatorial if and only if f is dictatorial, and $\hat{f}$ is manipulable (in the sense of Definition 4.3) if and only if f is manipulable (in the sense of Definition 5.2), and moreover, $\hat{f}$ is of a form to which Theorem 4.1 can be applied. To do so, we must show that $\Gamma(\mathscr{A}_k)$ is a linked top-2 domain, i.e., firstly, we must prove that the preferences in $\Gamma(\mathscr{A}_k)$ are top-2 preference relations, and secondly, knowing that $\Gamma(\mathscr{A}_k)$ thus is a top-2 domain, we must prove that $\Gamma(\mathscr{A}_k)$ is linked. Before we turn to the general proof, we will in an example, which then will serve as guideline for the general case, consider the instance when $M = 4$ and $k = 2$.

Example 5.6. Suppose that two of the alternatives in the set $\mathscr{A} = \{a_1, a_2, a_3, a_4\}$ have to be chosen, which means that one of the following six elements in $\mathscr{A}_2$ is going to be elected:

$$\{a_1, a_2\}, \quad \{a_1, a_3\}, \quad \{a_1, a_4\}, \quad \{a_2, a_3\}, \quad \{a_2, a_4\}, \quad \{a_3, a_4\} \tag{5.5}$$

First, we want to analyze the structure of preferences over $\mathscr{A}_2$ that are induced by preferences over $\mathscr{A}$. Consider therefore the two preferences $P, P' \in \Sigma$ defined by

$$\begin{aligned} & r_1(P) = a_1, \quad r_2(P) = a_2, \quad r_3(P) = a_3, \quad \text{and} \quad r_4(P) = a_4 \\ \text{and} \qquad & r_1(P') = a_1, \quad r_2(P') = a_3, \quad r_3(P') = a_2, \quad \text{and} \quad r_4(P') = a_4, \end{aligned}$$

and let $\hat{P}$ and $\hat{P}'$ be the partial preferences induced by P respectively P' on $\mathscr{A}_2$. Clearly, the most preferred alternative for $\hat{P}$ in $\mathscr{A}_2$ must be the subset consisting of P's two first alternatives, that is $r_1(\hat{P}) = \{a_1, a_2\}$. But a moment of thought should convince the reader that the set $\{a_1, a_3\}$ is a second best of alternative for $\hat{P}$, i.e., apart from $\{a_1, a_2\}$, the set $\{a_1, a_3\}$ is preferred to any other element in $\mathscr{A}_2$. Thus, we have $r_2(\hat{P}) = \{a_1, a_3\}$. Hence, $\hat{P}$ is a top-2 preference relation, and since the consideration above applies to all preferences in $\Gamma(\mathscr{A}_k)$, we conclude that $\Gamma(\mathscr{A}_2)$ is a top-2 domain.

Next, we can argue that $\Gamma(\mathscr{A}_2)$ is linked. The two top alternatives in $\mathscr{A}_2$ according to $\hat{P}$ respectively $\hat{P}'$ are

$$\begin{aligned} & r_1(\hat{P}) = \{a_1, a_2\}, \quad r_2(\hat{P}) = \{a_1, a_3\}, \\ \text{and} \qquad & r_1(\hat{P}') = \{a_1, a_3\}, \quad r_2(\hat{P}') = \{a_1, a_2\}, \end{aligned}$$

which means that $\{a_1, a_2\}$ and $\{a_1, a_3\}$ are connected in $\Gamma(\mathscr{A}_2)$. This indicates that two subsets in $\mathscr{A}_2$ are connected whenever they have exactly one element in common. Consider now the elements in $\mathscr{A}_2$ in the order in which they are listed in (5.5). The first two subsets are connected, and beginning with the third subset, every subset has one element in common with at least two of the preceding subsets and is hence connected to them. Thus, $\Gamma(\mathscr{A}_2)$ is a linked top-2 domain, and therefore, all strategy-proof social choice functions on $\Gamma(\mathscr{A}_2)$ that respect unanimity must by Theorem 4.1 be dictatorial. $\square$

We formalize now the steps in the example above for general values of k and M, namely, we will show that $\Gamma(\mathscr{A}_k)$ is a top-2 domain (Lemma 5.2), that two elements in $\mathscr{A}_k$ are connected if and only if they have exactly $k-1$ elements in common (Lemma 5.3), and finally, that $\Gamma(\mathscr{A}_k)$ is linked (Lemma 5.4).

Lemma 5.2. *The partial preference $\hat{P}$ on $\mathscr{A}_k$ that is induced by the preference $P \in \Sigma$ is a top-2 preference relation for $1 \le k \le M-1$, and*

$$r_1(\hat{P}) = \big\{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_k(P)\big\} \tag{5.6}$$

$$\textit{and} \qquad r_2(\hat{P}) = \big\{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_{k+1}(P)\big\}. \tag{5.7}$$

Proof. Firstly, in order to prove (5.6), set $A_1 = \big\{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_k(P)\big\}$. We have to show that $A_1 \hat{P} A$ for all $A \in \mathscr{A}_k \setminus \{A_1\}$. But once we have noted that if $A \in \mathscr{A}_k \setminus \{A_1\}$ and $a \in A \setminus A_1$, then a must be ranked below $r_k(P)$ by P, this is almost obvious, because every alternative in $A_1 \setminus A$ is preferred to every alternative in $A \setminus A_1$, and therefore $A_1 \hat{P} A$ according to Definition 5.1.

Secondly, for (5.7), we set $A_2 = \big\{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_{k+1}(P)\big\}$, and must show that $A_2 \hat{P} A$ for all $A \in \mathscr{A}_k \setminus \{A_1, A_2\}$. Fix therefore some $A \in \mathscr{A}_k \setminus \{A_1, A_2\}$, and suppose first that $r_k(P) \notin A$. Then all $a \in A \setminus A_2$ are ranked below $r_{k+1}(P)$ by

P, and the same argument as above shows that $A_2 \hat{P} A$. Next, if $r_k(P) \in A$, then there must be some j such that $1 \leq j \leq k-1$ and $r_j(P) \notin A$, because $A \neq A_1$. If A' denotes the set obtained by replacing $r_k(P)$ in A by $r_j(P)$, then clearly $A' \hat{P} A$. But since $r_k(P) \notin A'$, we have $A_2 \hat{P} A'$, and by transitivity therefore also $A_2 \hat{P} A$. □

Lemma 5.3. *For $1 \leq k \leq M-1$, two subsets $A_1, A_2 \in \mathscr{A}_k$ are connected if and only if they have exactly $k-1$ elements in common.*

Proof. Suppose first that A_1 and A_2 have exactly $k-1$ elements in common, that is, $A_1 = \{b_1, b_2, \ldots, b_{k-1}, b_k\}$ and $A_2 = \{b_1, b_2, \ldots, b_{k-1}, b'_k\}$ for some alternatives $b_1, b_2, \ldots, b_{k-1}, b_k, b'_k \in \mathscr{A}$. Let $P_1, P_2 \in \Sigma$ be some preferences with

$$\begin{aligned} & r_1(P_1) = b_1,\ r_2(P_1) = b_2, \ldots,\ r_{k-1}(P_1) = b_{k-1},\ r_k(P_1) = b_k,\ r_{k+1}(P_1) = b'_k \\ \text{and} \quad & r_1(P_2) = b_1,\ r_2(P_2) = b_2, \ldots,\ r_{k-1}(P_2) = b_{k-1},\ r_k(P_2) = b'_k,\ r_{k+1}(P_2) = b_k, \end{aligned}$$

Let $\hat{P}_1$ and $\hat{P}_2$ be the preferences on $\mathscr{A}_k$ that are induced by P_1 respectively P_2. According to Lemma 5.2, we have

$$\begin{aligned} & r_1(\hat{P}_1) = A_1, \quad r_2(\hat{P}_1) = A_2 \\ \text{and} \qquad & r_1(\hat{P}_2) = A_2, \quad r_2(\hat{P}_2) = A_1, \end{aligned}$$

which means that A_1 and A_2 are connected.

Conversely, if A_1 and A_2 are connected, then there exists a partial preference $\hat{P}$, induced by $P \in \Sigma$, such that $r_1(\hat{P}) = A_1$ and $r_2(\hat{P}) = A_2$. By Lemma 5.2, we have

$$\begin{aligned} & A_1 = \{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_k(P)\} \\ \text{and} \qquad & A_2 = \{r_1(P), r_2(P), \ldots, r_{k-1}(P), r_{k+1}(P)\}. \end{aligned}$$

Obviously, A_1 and A_2 have exactly $k-1$ elements in common. □

Lemma 5.4. *For $1 \leq k \leq M-1$, the domain $\Gamma(\mathscr{A}_k)$ is a linked top-2 domain.*

Proof. According to the definition of linked top-2 domains, we have to show that the elements in $\mathscr{A}_k$ can be ordered in a finite sequence $A_1, A_2, A_3, \ldots$ in such a way that A_1 and A_2 are connected, and from A_3 on, every subset A_i is connected to at least two of the preceding subsets. We claim that this requirement is met when the sets in $\mathscr{A}_k$ are ordered lexicographically with respect to a list $a_1, a_2, \ldots, a_M$ of the alternatives in $\mathscr{A}$; by this, we mean that a subset A_i is ordered before A_j if and only if the first alternative in $a_1, a_2, \ldots, a_M$ that belongs to only one of the two subsets belongs to A_i.[25] Ordered in this way, the first $M-k+1$ subsets (and only these) contain the

[25]For instance, the subsets in (5.5) are ordered lexicographically with respect to a_1, a_2, a_3, a_4.

common subset $\{a_1, a_2, \ldots, a_{k-1}\}$. By Lemma 5.3, any two of these subsets are connected, which in particular means that $A_1 \sim A_2$, and if $3 \leq i \leq M-k+1$, then A_i is connected with at least two of the preceding subsets. Next, if A_i is such that $i > M-k+1$, then there must be a first alternative a_l in $\{a_1, a_2, \ldots, a_{k-1}\}$ that does not belong to A_i, and since $\sharp A_i = k$, there must be at least two alternative a_j and $a_{j'}$ that belong to A_i, but not to $\{a_1, a_2, \ldots, a_{k-1}\}$. Replacing a_j respectively $a_{j'}$ by a_l, we obtain two sets in $\mathscr{A}_k$ that are ordered before A_i and which according to Lemma 5.3 are connected with A_i. Thus, $\Gamma(\mathscr{A}_k)$ is linked. □

It is now straightforward to see why Theorem 5.1 must be true: On the one hand, if a social choice function $f : \Sigma^N \to \mathscr{A}_k$ is dictatorial, then f is clearly strategy-proof because a dictator will definitely be worse off by misrepresentation, while the other individuals in the society have no influence on the social choice and can therefore not manipulate f. On the other hand, if $f : \Sigma^N \to \mathscr{A}_k$ is strategy-proof and satisfies complete unanimity, then the corresponding function $\hat{f} : \Gamma(\mathscr{A}_k)^N \to \mathscr{A}_k$ defined by (5.4) is strategy-proof and satisfies (ordinary) unanimity. Since $\Gamma(\mathscr{A}_k)$ according to Lemma 5.4 is a linked top-2 domain, we can apply Theorem 4.1 and conclude that $\hat{f}$, and therefore also f, must be dictatorial. By this, Question 2 from the introduction is answered.

5.3 The Case of Single-peaked Preferences

After having shown in the previous section that non-dictatorial strategy-proof social choice of fixed-sized subsets is impossible in general, we turn now to Question 3 from the introduction and investigate whether a possibility result can be obtained when voters' preferences are single-peaked. Assume therefore throughout this section that the set of alternatives $\mathscr{A}$ is equipped with a linear order $\prec$, and let $\Omega \subset \Sigma$ be a set of preferences that are single-peaked with respect to $\prec$. It is our purpose in the following to generalize the median rule from Section 3.2 to a social choice function for the social choice of fixed-sized subsets and to prove that it is strategy-proof. To this end, we will first show that the linear order $\prec$ on $\mathscr{A}$ and the single-peakedness of the preferences in Ω at least partly can be transferred to $\mathscr{A}_k$, namely when we restrict attention to those subsets that are *connected*; by this, we mean subsets $A \in \mathscr{A}_k$ with the property that whenever two alternatives a_i and a_j belong to A, then also all alternatives between a_i and a_j belong to A, or more formally:

Definition 5.3 (Connected Subset). A subset $A \subset \mathscr{A}$ is said to be *connected* if $a_i, a_j \in A$ and $a_i \prec a \prec a_j$ implies $a \in A$.

The set of all connected subsets of $\mathscr{A}$ containing exactly k elements is a proper subset of $\mathscr{A}_k$ and will be denoted by $\mathscr{A}_k^\circ$. The linear order $\prec$ on $\mathscr{A}$ induces in a natural way a linear order $\prec_k$ on $\mathscr{A}_k^\circ$. To see this, consider first the following six linearly ordered alternatives, where we have marked the five connected subsets that contain exactly two of the alternatives:

$$\underbrace{a_1 \prec a_2}_{A_1} \prec a_3 \ldots \quad a_1 \prec \overbrace{a_2 \prec a_3}^{A_2} \prec \overbrace{a_4 \prec a_5}^{A_4} \prec a_6, \quad \underbrace{a_3 \prec a_4}_{A_3}, \ \underbrace{a_5 \prec a_6}_{A_5}.$$

Here, it seems appropriate to say that A_i lies to the left of A_j if the left alternative in A_i lies to the left of the left alternative in A_j. For the general case, we formulate therefore the following definition.

Definition 5.4 (Induced Linear Order). If $\prec$ is a linear order on $\mathscr{A}$, we define the *induced linear order* $\prec_k$ on $\mathscr{A}_k^\circ$ in the following way: For $A_1, A_2 \in \mathscr{A}_k^\circ$, we will say that A_1 lies to the left of A_2, denoted by $A_1 \prec_k A_2$, if the leftmost alternative in A_1 lies to the left of the leftmost alternative in A_2.[26]

Our interest in connected subsets and the induced linear order is motivated by the following lemma, which states that if a voter has single-peaked preferences over the alternatives in $\mathscr{A}$, then his most preferred subset in $\mathscr{A}_k$ will be connected.

Lemma 5.5. *If $P \in \Omega$ and $\hat{P}$ is the preference induced by P on $\mathscr{A}_k$, then $r_1(\hat{P}) \in \mathscr{A}_k^\circ$.*

Proof. Suppose that $a_i, a_j \in r_1(\hat{P})$ and $a_i \prec a \prec a_j$, but $a \notin r_1(\hat{P})$. Obviously, $r_1(P) \in r_1(\hat{P})$, and thus either $a_i \prec a \prec r_1(P)$ or $r_1(P) \prec a \prec a_j$. If $a_i \prec a \prec r_1(P)$, then $a P a_i$, because P is single-peaked. But then we can obtain a set which is preferred to $r_1(\hat{P})$ by $\hat{P}$ if we replace a_i in $r_1(\hat{P})$ by a, which is absurd. A similar contradiction occurs if $r_1(P) \prec a \prec a_j$. Hence, we conclude that $a \in r_1(\hat{P})$. □

As a consequence of Lemma 5.5, the most preferred subsets of the voters in the society can be ordered linearly by $\prec_k$, and this enables us to transfer the concept of the median alternative to $\mathscr{A}_k^\circ$: In analogy with the definition of the median alternative in Section 3.2, we will say that a set $\bar{A} \in \mathscr{A}_k^\circ$ is a *median set* for the preference profile $\mathscr{P} \in \Omega^N$ if

$$\sharp \left\{ P_i \in \mathscr{P};\, r_1(\hat{P}_i) \succcurlyeq \bar{A} \right\} \geq \frac{N}{2}$$

$$\text{and} \qquad \sharp \left\{ P_i \in \mathscr{P};\, r_1(\hat{P}_i) \preccurlyeq \bar{A} \right\} \geq \frac{N}{2}.$$

[26] More formally, the induced linear order $\prec_k$ can be defined as follows: First, denote by $\min(A)$ the *minimal alternative* in A, which is that alternative $\bar{a} \in A$ that satisfies $\bar{a} \preccurlyeq a$ for all $a \in A$, and then we define $A_1 \prec_k A_2$ to hold if and only if $\min(A_1) \prec \min(A_2)$.

In order to generalize the median rule from Section 3.2, we want now to define a social choice function $f : \Omega^N \to \mathscr{A}_k$ that assigns to each preference profile its median set, but this function is well-defined only if we can ensure that a median set always exists and is unique, which is possible at least when N is odd:

Lemma 5.6. *If N is odd, every preference profile $\mathscr{P} \in \Omega^N$ has a unique median set.*

Proof. Let $A_1, A_2, \ldots, A_{M-k+1}$ be a list of the $M-k+1$ sets in $\mathscr{A}_k^\circ$, and define for $j = 1, 2, \ldots, M-k+1$ the cardinality $n_j = \sharp\,\{P_i \in \mathscr{P};\ r_1(\hat{P}_i) = A_j\}$. Note that $n_1 + n_2 + \ldots + n_{M-k+1} = N$, and moreover, $\sharp\,\{P_i \in \mathscr{P};\ r_1(\hat{P}_i) \preccurlyeq_k A_l\} = n_1 + n_2 + \ldots + n_l$ and $\sharp\,\{P_i \in \mathscr{P};\ r_1(\hat{P}_i) \succcurlyeq_k A_l\} = n_l + n_{l+1} + \ldots + n_{M-k+1}$. To prove the existence of a median set for a preference profile $\mathscr{P} \in \Omega^N$, let l be the *least* index such that $n_1 + n_2 + \ldots + n_l \geq N/2$. Then $n_1 + n_2 + \ldots + n_{l-1} < N/2$, and hence

$$n_l + n_{l+1} + \ldots + n_{M-k+1} = N - (n_1 + n_2 + \ldots + n_{l-1}) \geq \frac{N}{2}.$$

Thus, A_l is a median set for $\mathscr{P}$. To see that the median set for $\mathscr{P}$ is unique when N is odd, suppose that A_l and $A_{l'}$ are two different median sets for $\mathscr{P}$, and assume that $l < l'$. Note first that if N is odd, then a natural number that is larger than $N/2$ is larger than or equal to $(N+1)/2$. Since A_l and $A_{l'}$ are median sets, we have therefore both

$$n_1 + n_2 + \ldots + n_l \geq \frac{N+1}{2} \qquad \text{and} \qquad n_{l'} + n_{l'+1} + \ldots + n_{M-k+1} \geq \frac{N+1}{2}.$$

But then we obtain the contradiction

$$N = n_1 + n_2 + \ldots + n_{M-k+1} \geq n_1 + n_2 + \ldots + n_l + n_{l'} + n_{l'+1} + \ldots + n_{M-k+1} \geq N + 1,$$

and hence, the median set must be unique. □

For odd N, we define now the *median set rule* to be that social choice function $f : \Omega^N \to \mathscr{A}_k$ that assigns to a preference profile $\mathscr{P} \in \Omega^N$ the median set of $\mathscr{P}$. Obviously, the median set rule is non-dictatorial and satisfies complete unanimity. Moreover, the median set rule is also strategy-proof. The reason for this is that if $\hat{P}$ is the preference induced on $\mathscr{A}_k$ by a preference $P \in \Omega$, then on $\mathscr{A}_k^\circ$, which is the set of all possible outcomes of the median set rule, $\hat{P}$ has essentially the same structure as a single-peaked preference; this is made precise and proved in the following lemma.[27]

[27] When claiming that $\hat{P}$ has essentially the same structure as a single-peaked preference, we mean that $\hat{P}$ satisfies the two conditions in (5.8) and (5.9), by which single-peakedness was defined in

Lemma 5.7. *If $P \in \Omega$ and $\hat{P}$ is the preference induced by P on $\mathscr{A}_k$, then*

$$r_1(\hat{P}) \preccurlyeq_k A_1 \prec_k A_2 \quad \Longrightarrow \quad A_1 \, \hat{P} \, A_2 \tag{5.8}$$

$$\textit{and} \qquad r_1(\hat{P}) \succcurlyeq_k A_1 \succ_k A_2 \quad \Longrightarrow \quad A_1 \, \hat{P} \, A_2 \tag{5.9}$$

holds for all $A_1, A_2 \in \mathscr{A}_k^\circ$.

Proof. In order to prove (5.8), let $A_1, A_2 \in \mathscr{A}_k^\circ$ be such that $r_1(\hat{P}) \preccurlyeq_k A_1 \prec_k A_2$, and consider first two alternatives $a_i, a_j \in \mathscr{A}$ with $a_i \in A_1 \setminus A_2$ and $a_j \in A_2 \setminus A_1$. Obviously, as a direct consequence of (5.8), we have $a_i \prec a_j$. If now $a_i \preccurlyeq r_1(P)$, then $a_i \in r_1(\hat{P})$, and the rank of a_i according to P is therefore at most k, whereas a_j, by (5.8), must have a rank higher than k, and we conclude that $a_i \, P \, a_j$. On the other hand, if $r_1(P) \prec a_i$, then we also have $a_i \, P \, a_j$, which in this case follows from the assumption that P is single-peaked with respect to $\prec$. Hence, every alternative in $A_1 \setminus A_2$ is preferred to every alternative in $A_2 \setminus A_1$, which by Definition 5.1 means that $A_1 \, \hat{P} A_2$. This proves (5.8), and (5.9) can be proved in the same way. $\square$

Lemma 5.7 allows us to use the same arguments by which we proved the strategy-proofness of the median rule in Section 3.2 in order to prove that that the median set rule is strategy-proof:

Theorem 5.8. *The median set rule is strategy-proof.*

Proof. Let $\mathscr{P} \in \Omega^N$ be a preference profile, and let $\bar{A}$ be its median set. Consider an individual i. If $r_1(\hat{P}_i) = \bar{A}$, then individual i gets his top alternative and can therefore clearly not gain from misrepresentation. Consider therefore the case when $r_1(\hat{P}_i) \neq \bar{A}$, and assume, without loss of generality, that $r_1(\hat{P}_i) \succ \bar{A}$. Suppose now that individual i reports P_i' instead of his true preference P_i. If $r_1(\hat{P}_i') \succ \bar{A}$, then both

$$\sharp \left\{ P_j \in \mathscr{P};\, r_1(\hat{P}_j) \succcurlyeq \bar{A} \right\} \qquad \text{and} \qquad \sharp \left\{ P_j \in \mathscr{P};\, r_1(\hat{P}_j) \preccurlyeq \bar{A} \right\}$$

are unaffected, and therefore, $\bar{A}$ is also the median set of (P_i', P_{-i}). If, on the other hand, $r_1(\hat{P}_i') \preccurlyeq \bar{A}$, then

$$\begin{aligned} &\sharp \left\{ P_j \in (P_i', P_{-i});\, r_1(\hat{P}_j) \succcurlyeq \bar{A} \right\} \leq \sharp \left\{ P_j \in (P_i, P_{-i});\, r_1(\hat{P}_j) \succcurlyeq \bar{A} \right\} \\ \text{and} \qquad &\sharp \left\{ P_j \in (P_i', P_{-i});\, r_1(\hat{P}_j) \preccurlyeq \bar{A} \right\} \geq \sharp \left\{ P_j \in (P_i, P_{-i});\, r_1(\hat{P}_j) \preccurlyeq \bar{A} \right\}, \end{aligned}$$

Section 3.2. It is, however, worth noting that $\hat{P}$ is still incomplete on $\mathscr{A}_k^\circ$. To see this, let, for example, $a_1 \prec a_2 \prec a_3 \prec a_4$ be a linearly ordered set of alternatives, and define the preference P by $a_2 \, P \, a_3 \, P \, a_4 \, P \, a_1$. Then P is single-peaked with respect to $\prec$, but the preference $\hat{P}$ induced by P on $\mathscr{A}_2$ cannot rank the two connected subsets $\{a_1, a_2\}$ and $\{a_3, a_4\}$. Thus, $\hat{P}$ has in general less structure than a complete single-peaked preference, and $\hat{P}$ can, for instance, not be represented by a single-peaked utility function.

and hence, if $\bar{A}'$ is the median set of (P'_i, P_{-i}), then $\bar{A}' \preccurlyeq \bar{A}$. But since $\bar{A} \prec r_1(\hat{P}_i)$, we conclude by Lemma 5.7 that either $\bar{A} = \bar{A}'$ or $\bar{A}\hat{P}_i\bar{A}'$, and hence, individual i is not able to manipulate the median rule. The theorem is proved. $\square$

Recall that the median set rule is defined only when N is odd, but it is now straight forward to construct strategy-proof social choice functions $f : \Omega^N \to \mathscr{A}_k$ also for even N. Consider, for instance, that social choice function that assigns to the preference profile $(P_1, P_2, \ldots, P_N)$ the median of $(P_1, P_2, \ldots, P_N, P_N)$, where preference P_N has been doubled; if N is even, then this social choice function is well-defined, and using arguments similar to those in the proof of Theorem 5.8, one can show that it also is strategy-proof. Summarizing, we have thus shown that if voters' preferences are single-peaked, then there exist non-dictatorial strategy-proof social choice functions $f : \Omega^N \to \mathscr{A}_k$ that satisfy complete unanimity, and by this, Question 3 from the introduction is answered.

5.4 Related Literature

In this section, we discuss finally how the social choice of fixed-sized subsets studied in this thesis is related to other contributions in the literature of strategy-proof social choice theory. We mentioned already in the introduction that there, at least to our knowledge, exist no previous investigations of whether strategy-proof social choice of fixed-sized subsets is possible. However, there is a number of papers that study strategy-proof social choice of subsets of *variable* size, and they have their common starting point in the paper "Voting by committees" by Barberà et al. (1991). The original voting model in Barberà et al. (1991) is the following: A society consisting of N individuals has to choose a subset from a set $\mathscr{A}$ that contains M alternatives. Contrary to the social choice of fixed-sized subsets, however, there exist no restrictions on the number of alternatives that can be chosen, but any subset of $\mathscr{A}$ can be obtained as outcome, and the individuals are therefore assumed to have complete, antisymmetric and transitive preferences over the set of all subsets of $\mathscr{A}$.

In practice, voting situations of this type arise, for instance, in clubs that have to choose among candidates considered for membership. A concrete example is a choir to which a number of students apply, and a jury decides for each of the applicants after an admission test whether he or she is sufficiently talented to be accepted. Another instance where voting has the structure described in the model of Barberà et al. (1991) are decision-making committees, e.g., national parliaments, that consider a number of proposals and decide for each whether to accept it or to reject it.

In such voting situations, there is no direct conflict between the alternatives in the sense that the election of one alternative does not necessarily affect another alternative's possibility to be elected. This trait should be reflected in voters' preferences, and Barberà et al. (1991) assume therefore that voters are able to consider the alternatives in $\mathscr{A}$ one at a time and classify them either as *good*, meaning eligible, desirable, etc., or *bad*, then of course meaning ineligible, undesirable, etc. Formally, the set of voter i's good alternatives is denoted by G_i, and Barberà et al. (1991) assume that voters' preferences are *separable* in the sense that if A is a strict subset of $\mathscr{A}$ and $a \in \mathscr{A}$ is an alternative that does not belong to A, then adding a to A makes voter i better off if and only if he classifies a as good, or more formally

$$(A \cup \{a\}) \ P_i \ A \qquad \Longleftrightarrow \qquad a \in G_i. \tag{5.10}$$

Note that condition (5.10) is similar to condition (5.2) used in Section 5.1 in that both require preferences over subsets to be consistent with preferences over alternatives.[28]

Barberà et al. (1991) show that if voters' preferences over subsets of $\mathscr{A}$ are separable, then the social choice of subsets of $\mathscr{A}$ of variable size can be made strategy-proof, that is, if Ω denotes the set of all separable preferences over $2^{\mathscr{A}}$, then it is possible to find strategy-proof social choice functions of the form $f : \Omega^N \to 2^{\mathscr{A}}$, and in addition, Barberà et al. (1991) are able to characterize all strategy-proof social choice functions of the form $f : \Omega^N \to 2^{\mathscr{A}}$. The possibility result in Barberà et al. (1991) should not come as a surprise: Recall from Example 2.6 that when society has to choose one of two alternatives, then one can find strategy-proof social choice functions, e.g., the majority rule. The voting problem in Barberà et al. (1991) can be considered as repeated binary choice where society for each of the N alternatives in $\mathscr{A}$ must decide whether to include this alternative in the final subset or not, and since preferences are assumed to be separable, we should obtain a strategy-proof social choice function when we apply the majority rule to each alternative to decide whether to include it.

Comparing the voting model in Barberà et al. (1991) with the social choice of fixed-sized subsets, the following can be noted: In both voting models, society is going to choose a subset from a set $\mathscr{A}$, which, contrary to the single-valued choice in the Gibbard-Satterthwaite theorem, means that alternatives are not considered to be mutually exclusive. The fundamental difference between the two voting models is that in Barberà et al. (1991), society can choose any number of alternatives from

[28]In particular, this means that condition (5.10) also can fail to hold for reasons similar to those we discussed in Remark 5.1.

$\mathscr{A}$, whereas society in the voting model considered in this thesis must choose a fixed number of alternatives. This means that in Barberà et al. (1991), there is no conflict between the alternatives in $\mathscr{A}$ in the sense that the election of one alternative does not affect other alternatives' possibility to be elected, whereas in our approach, the alternatives are competing for the k places in the final subset. This difference leads to entirely different structures of individual preferences: In Barberà et al. (1991), the alternatives in $\mathscr{A}$ are evaluated *absolutely* in the sense that an alternative either is good or bad, disregarded from the qualities of other alternatives. This preference structure makes it possible to rank subsets of *different* sizes, which is formalized in (5.10). Note, however, that the condition in (5.10) has no implications on how subsets of the same size are ranked.[29] On the contrary, in the voting model in this thesis, we assume that voters have complete preferences over the alternatives in $\mathscr{A}$, which means that alternatives are evaluated *relatively* to each other in the sense that if a and b are two alternatives in $\mathscr{A}$, then either a is better than b, or b is better than a, but we have no information on whether a or b are desirable in an absolute sense. These preferences are then used to rank subsets of the *same* size via (5.2), whereas subsets of different sizes cannot be ranked by (5.2), which neither is necessary in our approach. Finally, it is important to note that the two voting models lead to fundamentally different results: while the social choice in the model of Barberà et al. (1991) can be made strategy-proof, this is not the case for the social choice of fixed-sized subsets.

The voting model in Barberà et al. (1991) has been modified and extended in different ways in a number of papers, e.g., in Serizawa (1995), Le Breton and Sen (1999), Aswal et al. (2003), Barberà et al. (2005), and Svensson and Torstensson (2005). Among these papers, Aswal et al. (2003) is of special interest for a comparison with the approach in this thesis. The basic voting model in this paper is of the same type as in Barberà et al. (1991), and voters' preferences are also assumed to be separable in the sense of (5.10), but in addition, Aswal et al. (2003) assume external restrictions on the number of elements in the subset of $\mathscr{A}$ which society is going to choose, which makes some elements in $2^{\mathscr{A}}$ infeasible. More

[29]One can of course object that the assumption in Barberà et al. (1991) that voters have complete preferences over the elements in $2^{\mathscr{A}}$ implies that if $a, b \in \mathscr{A}$, then $\{a\}$ and $\{b\}$ are comparable, and therefore it would not be unreasonable to assume also, in addition to the separability in (5.10), that if for instance $\{a\} P \{b\}$ and $A \subset \mathscr{A} \setminus \{a,b\}$, then $(A \cup \{a\}) P (A \cup \{b\})$. However, the separability restriction in (5.10) is already sufficient for the existence of non-dictatorial strategy-proof social choice functions, and additional restrictions will therefore not affect the possibility result in Barberà et al. (1991).

concretely, they consider restrictions that require the subset chosen by society to contain between k_1 and k_2 elements, where $0 < k_1 \leq k_2 < M$. An example of a situation that naturally gives rise to such restrictions is the following: An institute is in need of additional researchers for a new project, and advertises therefore a post. If there will be several very skilled applicants, the institute is willing to employ more than one researcher, but on the other hand, due to budget restrictions the institute will not be able to able to employ more than three researchers. Thus, we have $k_1 = 1$ and $k_2 = 3$ in this case.

Given the restriction $0 < k_1 \leq k_2 < M$, Aswal et al. (2003) prove that any strategy-proof social choice function has to be dictatorial, a result which thus is in sharp contrast to the result in Barberà et al. (1991). When $k_1 = k_2 = k$, we obtain as a special case of course the situation when society has to choose an element from $\mathscr{A}_k$, which seems to be quite similar to the social choice of fixed-sized subsets. Note, however, that the elements in $\mathscr{A}_k$ are not affected by the separability condition in (5.10) since they are all of the same size, which means that this special case of the result of Aswal et al. (2003) is based on the assumption that individual preferences over $\mathscr{A}_k$ are complete and unrestricted. In this light, the dictatorial result of Aswal et al. (2003) for the case when $k_1 = k_2$ is not surprising because it is a direct consequence of the original Gibbard-Satterthwaite theorem. On the other hand, condition (5.2) used in this thesis puts restrictions on voters' preferences over subsets of the same size, and therefore, our result on the non-existence of non-dictatorial strategy-proof social choice function choosing an element from $\mathscr{A}_k$ can *not* be seen as a special chase of the result in Aswal et al. (2003).

6 Conclusions

IT IS NOW time to summarize the results of our study of the strategy-proof social choice of fixed-sized subsets and to link up with the questions proposed in the introduction. We started our analysis with the observation that it is not obvious when a social choice function of the form $f : \Sigma^N \to \mathscr{A}_k$ should be said to be manipulable, but one needs to know how preferences over the alternatives in $\mathscr{A}$ should be translated to preferences over $\mathscr{A}_k$, which led us to Question 1. We answered this question by arguing that an individual prefers one subset to another if and only if the former can be obtained from the latter by successive replacement of worse alternatives by better alternatives.

Knowing the structure of preference over $\mathscr{A}_k$, we could then turn to the question whether strategy-proof social choice of fixed-sized subsets is possible, and first, we considered the case when the set of preferences over $\mathscr{A}$ is unrestricted. It turned out that this set of preferences induced a set of preferences on $\mathscr{A}_k$ that has a complicated structure, and in order to analyze this structure, we derived a general theoretical result, Theorem 4.1. This theorem serves thus as a tool in this thesis, but it is also of interest on its own: Firstly, it shows that a large class of restricted preference domains is dictatorial, and it can thus also be applied in other contexts. Secondly, it shows that the assumption of complete preferences in the Gibbard-Satterthwaite theorem can be relaxed considerably because it suffices to assume that preferences belong to a linked top-2 domain in order to conclude that every strategy-proof social choice function that satisfies unanimity must be dictatorial. Thirdly, and finally, it allows us to obtain a strengthening of the Gibbard-Satterthwaite theorem, which states that every non-dictatorial social choice function that satisfies unanimity is not only manipulable, but it can be manipulated in such a way that some individual obtains at least his second best alternative. This is our answer of Question 4, which of course is unrelated with the purpose of this thesis, but it is nevertheless of interest.

Applying Theorem 4.1 to the set of all preferences that are induced on $\mathscr{A}_k$, we could conclude that the social choice of fixed-sized subsets, precisely as the single-valued social choice in the original Gibbard-Satterthwaite theorem, cannot be made strategy-proof, which answers Question 2.

Finally, we considered also the case when voters' preferences over the alternatives in $\mathscr{A}$ are single-peaked since this is a reasonable assumption in many voting situations. As answer to Question 3, we found, again in analogy with the single-valued case, that single-peaked preferences admit non-dictatorial strategy-proof social choice functions for the social choice of fixed-sized subsets.

A Mathematical Appendix

IN THIS APPENDIX, we provide a short overview over the mathematical concepts used in the formalizations and proofs throughout this thesis. Even though the results in this thesis must be regarded non-trivial, their proofs do not require any advanced techniques from higher mathematics, but the only ingredients needed are the language of set theory and some fundamental methods for mathematical deduction. The parts of set theory needed for our purposes are presented in Section A.1. Among the methods for mathematical deduction used in our proofs, there are two that can be confusing to the reader unfamiliar with them, whence we will explain them here in some detail. These are proof by contradiction, considered in Section A.2, and the principle of mathematical induction, explained in Section A.3. The material presented here can, for instance, also be found in Sydsæter and Hammond (2006), or in the first part in Grimaldi (1998).

A.1 Elementary Set Theory

At several places in this thesis, we encounter a number of objects that are suitably considered as a whole. These may be the individuals that form society, the available alternatives among which society is going to choose, or the preferences that are reasonable in some context. Such a collection of objects is called a *set*, and is usually denoted by capitals as A, $\mathscr{A}$, $\mathcal{I}$, or Σ. The objects itself are usually denoted by lower-case letters, and if an object a belongs to a set A, we say that a is an *element* in A, and write $a \in A$. If a does not belong to A, we write $a \notin A$.

Sets can be described explicitly by listing their elements, enclosed by braces. For example, if a set $\mathscr{A}_1$ consists of the four alternatives a_1, a_2, a_3, and a_4, we write

$$\mathscr{A}_1 = \{a_1, a_2, a_3, a_4\}. \tag{A.1}$$

Alternatively, a set can be described indirectly by declaring a property that is distinguishing for its elements. E.g., suppose that the preference P orders the alternatives in the set $\mathscr{A}_1$ according to $a_1 \, P \, a_2 \, P \, a_3 \, P \, a_4$. Then we can define a new set $\mathscr{A}_2$ by requiring, for instance, that $\mathscr{A}_2$ should contain precisely those alternatives in $\mathscr{A}_1$ that are preferred to a_3 by P. For this, we will use the notation

$$\mathscr{A}_2 = \{a \in \mathscr{A}_1;\ a \, P \, a_3\}, \tag{A.2}$$

which should be read as "$\mathscr{A}_2$ is the set of all $a \in \mathscr{A}_1$ such that $a \, P \, a_3$". Of course, the set $\mathscr{A}_2$ consists of the two elements a_1 and a_2, so instead of the more complicated expression in (A.2) we could simpler write $\mathscr{A}_2 = \{a_1, a_2\}$. Often, however, it is more convenient to describe sets as in (A.2) because this way can reveal more of the structure of a set than a simple list can do. Also, sometimes it is not even possible to list the elements in a set, which, for instance, is the case when the set contains infinitely many elements.

Throughout this thesis, we will only consider sets that contain finitely many elements, and the number of elements in a set A will be denoted by $\sharp A$; for example, for the set $\mathscr{A}_1$ in (A.1), we have $\sharp \mathscr{A}_1 = 4$.

Given two sets A and B, we say that A is a *subset* of B, denoted by $A \subset B$, if every element in A also is an element in B. For instance, the set $\mathscr{A}_2$ in (A.2) is a subset of the set $\mathscr{A}_1$ in (A.1). For a set A, the set of all subsets of A is called the *power set* of A and is denoted by 2^A. For example, if $\mathscr{A}_3 = \{a_1, a_2, a_3\}$, then

$$2^{\mathscr{A}_3} = \Big\{\emptyset, \{a_1\}, \{a_2\}, \{a_3\}, \{a_1, a_2\}, \{a_1, a_3\}, \{a_2, a_3\}, \{a_1, a_2, a_3\}\Big\},$$

where $\emptyset$ denotes the *empty set*, i.e., the set that does not contain any elements at all.

Starting with two sets A and B, we can construct a number of new sets by combining the elements in A and B in different ways, and in this thesis, we need the following concepts:

1. The *union* of A and B, denoted by $A \cup B$, contains those elements that either belong to A, to B, or to both, i.e., $A \cup B = \{a;\ a \in A \text{ or } a \in B\}$.

2. The *intersection* of A and B, denoted by $A \cap B$, contains those elements that belong to both A and B, i.e., $A \cap B = \{a;\ a \in A \text{ and } a \in B\}$.

3. The *difference* of A and B, denoted by $A \setminus B$, contains those elements that belong to A, but not to B, i.e., $A \setminus B = \{a;\ a \in A \text{ and } a \notin B\}$.

4. The *Cartesian product* of A and B, denoted by $A \times B$, consists of all ordered pairs whose first element belongs to A and whose second element belongs to B, i.e., $A \times B = \{(a, b);\ a \in A \text{ and } b \in B\}$.

5. The nth *power* of the set A, denoted by A^n, contains all n-tuples of elements in A, i.e., $A^n = \{(a_1, a_2, \ldots, a_n);\ a_1, a_2, \ldots, a_n \in A\}$.

Often, one does not study sets on their own, but is interested in how the elements in two sets are related. A special case of a relation between the elements in two sets A and B is a *function*, which is a rule that assigns to every element $a \in A$ one, and only one, element $b \in B$. When f is a function from A to B, this will be indicated by $f : A \to B$, and if b is the element assigned to a by f, we write $f(a) = b$ and we will also say that a is the *argument* of f.

A.2 Proof by Contradiction

In many of the proofs in this thesis, we use formulations like “the statement will be proved by a contradiction argument” or “we will argue by contradiction”. By this, we indicate that the following procedure for mathematical deduction, known as *proof by contradiction*, will be used: Suppose that $\mathcal{S}$ is a statement we want to prove, and let $\mathcal{S}'$ be its logical contrary, that is, $\mathcal{S}'$ is true if and only if $\mathcal{S}$ is false.[30] In order to prove $\mathcal{S}$, which might be hard to do directly, we suppose for a while that $\mathcal{S}'$ would be true, and analyze the consequences of this assumption. More concretely, assuming that $\mathcal{S}'$ is true, we derive other statements that then also must be true. This will finally lead to a statement that is definitely false, from which we then can conclude that the initial assumption that $\mathcal{S}'$ holds was erroneous. But since exactly one of $\mathcal{S}$ and $\mathcal{S}'$ must be true, we have thus proved that $\mathcal{S}$ must hold.

We illustrate the method of proof by contradiction by a classical example. Recall first that a natural number greater than 1 is called a *prime number* if it is divisible only by 1 and itself. A natural number greater than 1 that is not a prime number can be written as a product of prime numbers, and is therefore said to be *composite*. Suppose now we want to prove the following statement:

There are infinitely many primes. (A.3)

It seems difficult to prove this statement directly, whence proof by contradiction is an appropriate method. Consider therefore the logical contrary of (A.3):

There is a finite number of primes. (A.4)

We need to show that (A.4) is not logically tenable. If (A.4) is true and the number of primes is N, say, then we can list all primes in a finite sequence as

$$p_1,\ p_2,\ \ldots,\ p_N. \tag{A.5}$$

Consider now the number p obtained by multiplying all primes and adding 1, i.e.,

$$p = p_1 \cdot p_2 \cdot \ldots \cdot p_N + 1.$$

Obviously, p cannot be equal to any of the prime numbers $p_1, p_2, \ldots, p_N$, and therefore, p must be composite. On the other hand, p cannot be divisible by p_1, because dividing p by p_1 leaves a remainder of 1. For the same reason, p cannot be divisible

[30] For instance, if $\mathcal{S}$ denotes the statement “all swans are white”, then $\mathcal{S}'$ stands for “there is at least one swan that is not white”.

by p_2, by p_3, or by any other of the prime numbers in (A.5). But then, p cannot be composite, because a composite number can be factored as a product of primes numbers. We are thus forced to conclude that p must be composite at the same time as p cannot be composite, which is obviously a contradiction; therefore, the statement in (A.4), from which this contradiction was derived, must be wrong, which implies that (A.3) must be true, and the proof is fulfilled.

A.3 The Principle of Mathematical Induction

Many theorems in social choice theory contain one or more parameters. Consider, for instance, the fundamental result in strategy-proof social choice theory:

> ***The Gibbard-Satterthwaite theorem:*** *Let $\mathscr{A}$ be a set containing $M \geq 3$ elements, and denote by Σ the set of all strict preferences over $\mathscr{A}$. Let further $f : \Sigma^N \to \mathscr{A}$ be a social choice function that satisfies unanimity. Then f is strategy-proof if and only if f is dictatorial.*

This statement contains two parameters—firstly, M, the number of alternatives in $\mathscr{A}$, and secondly, N, the number of arguments in f—and the theorem is claimed to be true for all $M \in \{3,4,5,\ldots\}$ and all $N \in \{2,3,4,\ldots\}$.[31] There exist different strategies to prove statements involving parameters. Sometimes, one is able to construct an argument that is independent of the actual value of the parameter; for instance, in Appendix B we prove the Gibbard-Satterthwaite theorem for the case $N = 2$ in a way which does not depend on M, provided that $M \geq 3$, which means that we actually prove the theorem for all $M \in \{3,4,5,\ldots\}$.

Often, however, it is not possible to succeed in such a simple way, but one is in need of more sophisticated strategies. A standard technique for proving theorems that contain parameters that are natural numbers[32] is known as *proof by induction*. Thereby, one splits the proof of a certain statement that depends on a parameter N that is a natural number into the following two steps:

1. Show that the statement holds for $N = 1$.

2. Show that if the statement is true for a particular natural number, say $\bar{N}$, then it is also true for the proceeding natural number, that is, for $\bar{N}+1$.[33]

[31]This theorem is of course also true for $N = 1$, but it is interesting only for $N \geq 2$.

[32]By *natural numbers*, we refer here to the set $\{1,2,3,\ldots\}$, which is also the convention used in Sydsæter and Hammond (2006). In many mathematical contexts, however, the term “natural numbers” is reserved for the set $\{0,1,2,3,\ldots\}$.

[33]This step is sometimes called the *induction step*, and its premise that the theorem is true for a particular $\bar{N}$ is called the *induction hypothesis*.

In combination, these two steps imply that the statement holds for all natural numbers $N \in \{1, 2, 3, \ldots\}$.[34]

In this thesis, induction proofs are used at several places, for instance, when we show that the Gibbard-Satterthwaite theorem holds for all $N \geq 2$. In order to make the reader familiar with the structure of a proof by induction, we illustrate this technique here by an elementary example. Suppose we want to prove the summation formula

$$1 + 2 + 3 + \cdots + N = \frac{N(N+1)}{2}, \tag{A.6}$$

which is claimed to be true for all $N \geq 1$. For a clearer argumentation, we denote the left hand side by $f(N)$ and the right hand side by $g(N)$, i.e.,

$$f(N) = 1 + 2 + 3 + \cdots + N \qquad \text{and} \qquad g(N) = \frac{N(N+1)}{2}.$$

With this notation, we have to show that

$$f(N) = g(N) \tag{A.7}$$

holds for all $N \geq 1$, which will be done by carrying out the two steps required for an induction proof.

First, if $N = 1$, then $f(1) = 1$ and $g(1) = 1 \cdot (1+1)/2 = 1$, which shows that (A.7) is true for $N = 1$.

Second, we have to carry out the induction step by showing that *if* $f(\bar{N}) = g(\bar{N})$ for a particular $\bar{N}$, then also $f(\bar{N}+1) = g(\bar{N}+1)$. But if $f(\bar{N}) = g(\bar{N})$ for some $\bar{N}$, then straightforward calculations give

$$f(\bar{N}+1) = 1 + 2 + \ldots + \bar{N} + (\bar{N}+1) = f(\bar{N}) + (\bar{N}+1) = g(\bar{N}) + (\bar{N}+1) =$$
$$= \frac{\bar{N}(\bar{N}+1)}{2} + (\bar{N}+1) = \frac{\bar{N}(\bar{N}+1) + 2(\bar{N}+1)}{2} = \frac{(\bar{N}+1)(\bar{N}+2)}{2} = g(\bar{N}+1),$$

which is exactly what we needed to show. Note that we at the third equal sign use the induction hypothesis that $f(\bar{N}) = g(\bar{N})$.

[34] This principle works like a domino effect: If dominos have been placed standing in an infinite ray, the first domino is knocked over, and every domino that is knocked over also knocks the following domino over, then the principle of mathematical induction says that every domino sooner or later is knocked over.

B A Proof of the Gibbard-Satterthwaite Theorem

THE PURPOSE OF this appendix is twofold: firstly, we present a complete proof of the Gibbard-Satterthwaite theorem (Theorem 2.3 in this thesis), the main result in strategy-proof social choice theory, and secondly, we show why the proof of Theorem 3.2 given in Aswal et al. (2003) cannot be applied to partial preference relations, thereby motivating our proof of Theorem 4.1 given in Section 4.2.

We will prove the Gibbard-Satterthwaite theorem using induction over the number of individuals: First, we show in Lemma B.1 that the theorem holds for $N = 2$ individuals, and then, we prove in Lemma B.2 that if the theorem holds for N individuals, then it is also true for $N+1$ individuals. In combination, these two steps provide a complete proof of the Gibbard-Satterthwaite theorem.

Lemma B.1. *Suppose that $\mathscr{A}$ is a set containing at least three alternatives, and let $f : \Sigma^2 \to \mathscr{A}$ be a social choice function that satisfies unanimity. Then f is strategy-proof if and only if f is dictatorial.*

The following proof follows closely a corresponding proof in Svensson (1999).

Proof. On the one hand, if f is dictatorial, the dictator is of course best off if he presents his preference truly, whereas the other individual in the society cannot affect the social choice and hence is neither able to gain from misrepresentation. Thus, f is strategy-proof in this case.

On the other hand, suppose now that f is strategy-proof, and recall from Section 2.2 that a strategy-proof social choice function satisfies monotonicity and Pareto optimality. Let a_1 and a_2 be two alternatives in $\mathscr{A}$, and consider the following preferences:

P_1	P_2	P_2'
a_1	a_2	a_2
a_2	a_1	$\vdots$
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	a_1

Due to Pareto optimality we must have either $f(P_1,P_2) = a_1$ or $f(P_1,P_2) = a_2$, and suppose for simplicity that $f(P_1,P_2) = a_1$. Consider now preference P_2' where a_1 is ranked last, and note that Pareto optimality also here implies $f(P_1,P_2') \in \{a_1,a_2\}$. However, if $f(P_1,P_2') = a_2$, then individual 2 would be able to manipulate f by going from (P_1,P_2) to (P_1,P_2'), and hence $f(P_1,P_2') = a_1$. But monotonicity implies then that $f(P_1,P_2) = a_1$ whenever $r_1(P_1) = a_1$, that is, individual 1 is a dictator for alternative a_1.

In order to show that individual 1 now must be a dictator also for all other alternatives, suppose that $a \in \mathscr{A} \setminus \{a_1, a_2\}$, and consider the following two preference profiles

$$\begin{array}{cc} P_1 & P_2 \\ \hline a & a_2 \\ a_2 & a \\ \vdots & \vdots \end{array} \qquad \text{and} \qquad \begin{array}{cc} P_1 & P_2 \\ \hline a_2 & a \\ a & a_2 \\ \vdots & \vdots \end{array}$$

Applying the same arguments as in the first paragraph to the left preference profile, we conclude that either individual 1 is a dictator for a or individual 2 is a dictator for a_2, but the first paragraph shows also that the latter case is not possible. Hence, individual 1 is a dictator for all alternatives in $\mathscr{A}$, possibly with the exception of a_2. But if individual 1 is not a dictator for a_2, then the right preference profile and an argument as in the first paragraph show that individual 2 must be a dictator for a, which is not possible because $P_1, P_2 \in \Sigma$, $r_1(P_1) = a_1$ and $r_1(P_2) = a$ would then imply both $f(P_1, P_2) = a_1$ and $f(P_1, P_2) = a$. Thus, individual 1 must be a dictator for f. $\square$

Lemma B.2 below provides the induction step not only for the original Gibbard-Satterthwaite theorem, where the domain of preferences is assumed to be unrestricted, but for a large class of restricted domains, namely all domains that are minimally rich[35], and it is therefore of interest on its own. It has also been used by Aswal et al. (2003) as part of their proof of Theorem 3.2. The formulation of Lemma B.2 given below is essentially the same as in Aswal et al. (2003), but the proof follows closely that given in Svensson (1999), because it is simpler than in Aswal et al. (2003). We would also like to remark that results similar to Lemma B.2 can be found, e.g., in Kalai and Muller (1977) (with another notion of manipulability than in this thesis), and Barberà and Peleg (1990) (under the assumption of unrestricted preferences).

Lemma B.2. *Suppose that $\Omega \subset \Sigma$ is a minimally rich domain over the set $\mathscr{A}$ of alternatives. If the implication*

$$\left.\begin{array}{c} \textit{a social choice function } f : \Omega^n \to \mathscr{A} \textit{ is} \\ \textit{strategy-proof and satisfies unanimity} \end{array}\right\} \implies f \textit{ is dictatorial} \qquad \text{(B.1)}$$

holds for all n such that $2 \le n \le N$, then it also holds for $n = N + 1$.

[35]A domain Ω is said to be *minimally rich* if for every $a \in \mathscr{A}$ there exists some $P \in \Omega$ such that $r_1(P) = a$. Note, for example, that single-peaked domains, linked domains and the domain of unrestricted preferences are minimally rich.

Proof. Suppose that $f : \Omega^{N+1} \to \mathscr{A}$ is a strategy-proof social choice function that satisfies unanimity, and define the social choice function $g : \Omega^2 \to \mathscr{A}$ by

$$g(P_1, P_2) = f(P_1, \underbrace{P_2, \ldots, P_2}_{N \text{ copies}}).$$

The aim in the first part of this proof is to show that g is dictatorial. We note first that g satisfies unanimity because f satisfies unanimity. Next, we claim that g is strategy-proof. It is clear that g is strategy-proof in its first argument because if individual 1 can manipulate g at P_1, then individual 1 can of course also manipulate f at P_1. To show that g is strategy-proof in its second argument, we argue by contradiction. Suppose therefore that

$$g(P_1, P_2) = a, \quad g(P_1, P_2') = b, \quad \text{and} \quad b\, P_2\, a \tag{B.2}$$

for some preferences $P_1, P_2, P_2' \in \Omega$. Define

$$c_k = f(P_1, \underbrace{P_2', \ldots, P_2'}_{k \text{ copies}}, \underbrace{P_2, \ldots, P_2}_{N-k \text{ copies}}),$$

and note that $c_0 = a$ and $c_N = b$. Comparing c_k with c_{k+1}, we observe that if

$$c_{k+1}\, P_2\, c_k\,, \tag{B.3}$$

then individual $k+2$ can manipulate f by representing P_2' instead of P_2. Since the preferences in Ω are complete and f is strategy-proof, we conclude therefore that

$$c_{k+1} = c_k \quad \text{or} \quad c_k\, P_2\, c_{k+1}\,. \tag{B.4}$$

Considering (B.4) for all $k = 0, 1, \ldots, N$ and using transitivity, we obtain either $a = c_0 = c_N = b$, which is absurd since $a \neq b$, or $a P_2 b$, which contradicts the assumptions in (B.2). Thus, g must be strategy-proof. Now, we can apply the implication in (B.1) to g, and conclude that g must be dictatorial. This means that either individual 1 or individual 2 is a dictator for g, and in the rest of this proof we show that both cases imply that also f is dictatorial.

First, if individual 1 is the dictator for g, then we can use the monotonicity property of strategy-proof social choice functions (Lemma 2.4) and an argument similar to that in the second part of the proof of Theorem 4.1 on page 43, in order to conclude that individual 1 also is a dictator for f.

Second, consider the case when individual 2 is the dictator for g. Let $\bar{P}_1 \in \Omega$ be some fixed preference, and define the social choice function $h : \Omega^N \to \mathscr{A}$ by

$$h(P_2, P_3, \ldots, P_{N+1}) = f(\bar{P}_1, P_2, P_3, \ldots, P_{N+1}).$$

We note that h must be strategy-proof because f is strategy-proof. Moreover, since

$$h(P_2, P_2, \ldots, P_2) = g(\bar{P}_1, P_2) = r_1(P_2),$$

we conclude, using monotonicity, that h satisfies unanimity. Thus, we can apply the implication in (B.1) in order to conclude that h is dictatorial. Assume, without loss of generality, that individual 2 is the dictator for h, that is

$$h(P_2, P_3, \ldots, P_{N+1}) = r_1(P_2).$$

Next, let $\bar{P}_3, \ldots, \bar{P}_{N+1} \in \Omega$ be $N-1$ fixed preferences, and consider the social choice function $q : \Omega^2 \to \mathscr{A}$ defined by

$$q(P_1, P_2) = f(P_1, P_2, \bar{P}_3, \ldots, \bar{P}_{N+1}).$$

Observe that q is strategy-proof, because f is strategy-proof, and since

$$q(\bar{P}_1, P_2) = h(P_2, \bar{P}_3, \ldots, \bar{P}_{N+1}) = r_1(P_2) \tag{B.5}$$

we can apply monotonicity to $\bar{P}_1$, and conclude that q satisfies unanimity. Applying then the implication in (B.1) to q, we conclude that q is dictatorial, and (B.5) implies that individual 2 must be the dictator for q. Therefore, we have

$$f(P_1, P_2, \bar{P}_3, \ldots, \bar{P}_{N+1}) = q(P_1, P_2) = r_1(P_2). \tag{B.6}$$

Since equation (B.6) holds independently of the particular choice of the preferences $\bar{P}_3, \ldots, \bar{P}_{N+1}$, we conclude that individual 2 in fact is a dictator for f. The lemma is proved. □

We can now explain the complication that came up when we tried to transfer the proof of Theorem 3.2 to linked top-2 domains. As mentioned above, Lemma B.2 provides the induction step in the proof of Theorem 3.2 given in Aswal et al. (2003), which works because linked domains are minimally rich. However, the proof of Lemma B.2 contains an implication that does not hold for partial preference relations, namely, the passage from (B.3) to (B.4): if P_2 is a strict partial preference relation and we know that $c_{k+1} P_2 c_k$ does not hold, then there are three possibilities left (and not only two as in (B.4)), namely, either we have $c_{k+1} = c_k$, or $c_k P_2 c_{k+1}$, or c_k and c_{k+1} are not ranked by P_2, and in the latter case we are not able to derive the contradiction needed to show that g is strategy-proof. There seems to be no simple way to solve this complication, and therefore, we were forced to modify the proof of Theorem 3.2 in a more fundamental way. Whether Lemma B.2 actually holds for minimally rich partial preference domains remains an open research question.

References

Arrow, Kenneth (1963). *Social Choice and Individual Values*. Yale University Press, second edition.

Aswal, Navin, Shurojit Chatterji, and Arunava Sen (2003). Dictatorial domains. *Economic Theory* 22, 45–62.

Barberà, Salvador (2001). An introduction to strategy-proof social choice functions. *Social Choice and Welfare* 18: 619–653.

Barberà, Salvador, Jordi Massó, and Alejandro Neme (2005). Voting by committees under constraints. *Journal of Economic Theory* 122: 185–205.

Barberà, Salvador and Bezalel Peleg (1990). Strategy-proof voting schemes with continuous preferences. *Social Choice and Welfare* 7: 31–38.

Barberà, Salvador, Hugo Sonnenschein and Lin Zhou (1991). Voting by committees. *Econometrica*, Vol. 59, No. 3, 595–609.

Black, Duncan (1948). On the rationale of group decision-making. *Journal of Political Economy* 56, 23–34.

Bossert, Walter and John Weymark (2006). Social choice: Recent developments. *Working Paper No. 06-W03*. Department of Economics, Vanderbilt University, Nashville.

Gaertner, Wulf (2001). *Domain Conditions in Social Choice Theory*. Cambridge University Press.

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

Grimaldi, Ralph (1998). *Discrete and Combinatorial Mathematics: An Applied Introduction*. Addison Wesley, fourth edition.

Kalai, Ehud and Eitan Muller (1977). Characterization of domains admitting nondictatorial social welfare functions and nonmanipulable voting procedures. *Journal of Economic Theory* 16, 457–469.

Le Breton, Michel and Arunava Sen (1999). Separable preferences, strategyproofness, and decomposability. *Econometrica*, Vol. 67, No. 3, pp. 605–628.

Mas-Colell, Andreu, Michael Whinston, and Jerry Green (1995). *Microeconomic Theory*. Oxford University Press.

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

Sen, Amartya (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

Sen, Amartya (1999). The possibility of social choice. *The American Economic Review*, Vol. 89, No. 3, pp. 349–378.

Serizawa, Shigehiro (1995). Power of voters and domain of preferences where voting by committees is strategy-proof. *Journal of Economic Theory*, 67, 599–608.

Sprumont, Yves (1995). Strategyproof collective choice in economic and political environments. *The Canadian Journal of Economics*, Vol. 28, No. 1, 68–107.

Svensson, Lars-Gunnar (1999). The proof of the Gibbard-Satterthwaite theorem revisited. *Working Paper Series* 1999:1, Department of Economics, School of Economics and Management, University of Lund.

Svensson, Lars-Gunnar and Pär Torstensson (2005). Strategy-proof Allocation of Multiple Public Goods. *Working Paper Series* 2005:3, Department of Economics, School of Economics and Management, University of Lund.

Sydsæter, Knut and Peter Hammond (2006). *Essential Mathematics for Economic Analysis*. Prentice Hall, second edition.

Szpilrajn, Edward (1930). Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, Vol. 16, 386–389.

List of Notations

The page number indicates the page where the notation is introduced.

Basic Notations

$\mathscr{A}$	the set of alternatives, *8*
$\mathcal{I}$	the set of individuals, *8*
$\mathscr{A}_k$	the set of all subsets of $\mathscr{A}$ that contain exactly k elements, *45*
$\mathscr{A}_k^\circ$	the set of all connected subsets of $\mathscr{A}$ that contain exactly k elements, *57*

Preference Relations

aPb	alternative a is preferred to b according to preference P, *8*
$r_k(P)$	the kth ranked alternative of preference P, *9*
P_i	the preference of individual i, *9*
$\mathscr{P}$	a preference profile, i.e., $\mathscr{P} = (P_1, P_2, \ldots, P_N)$, *9*
P_{-i}	the preference profile of all individuals apart from individual i, *9*
$\mathscr{O}_{-i}(P_i, f)$	the option set with respect to P_i and f, *35*
$\hat{P}$	the preference induced by P on $\mathscr{A}_k$, *47*

Preference Domains

Σ	the domain of unrestricted preferences, *9*
Ω	a restricted preference domain, *19*
Γ	a top-2 domain, *29*
Γ_k	the set of all $P \in \Gamma$ such that $r_1(P) = a_k$, *35*
$\Gamma(\mathscr{A}_k)$	the set of all induced preferences over $\mathscr{A}_k$, *47*
$\prec$ and $\preccurlyeq$	an underlying linear order, *20*

Linked Domains

$a_i \sim a_j$	the alternatives a_i and a_j are connected, *23, 29*
$[a_{i_1}, a_{i_2}, \ldots, a_{i_n}]$	a chain connecting a_{i_1} with a_{i_n}, *30*

Set-theoretical Notations

$a \in A$	alternative a belongs to the set $\mathscr{A}$, *65*
$a \notin A$	alternative a does not belong to the set $\mathscr{A}$, *65*
$\sharp A$	the number of elements in the set $\mathscr{A}$, *66*
$A \cup B$	the union of the sets A and B, *66*
$A \cap B$	the intersection of the sets A and B, *66*
$A \setminus B$	the difference of the sets A and B, *66*
$A \times B$	the Cartesian product of A and B, *66*
$2^{\mathscr{A}}$	the power set of $\mathscr{A}$, *66*

Index