Single-Peaked Preferences

Extensions, Empirics and Experimental Results

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
(Doctor rerum politicarum)

von der Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte
DISSERTATION

von

Dipl. math. oec. Veronica Iris Block de Priego

Datum der mündlichen Prüfung:	06.02.2014
Referent:	Prof. Dr. Clemens Puppe
Korreferenten:	Prof. Dr. Johannes Philipp Reiß
	Prof. Dr. Sebastian Kube

Erklärung

gemäß §4, Abs. 4 der Promotionsordnung vom 15. August 2006

Ich versichere wahrheitsgemäß, die Dissertation bis auf die in der Abhandlung angegebene Hilfe selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und genau kenntlich gemacht zu haben, was aus Arbeiten anderer und aus eigenen Veröffentlichungen unverändert oder mit Abänderungen entnommen wurde.

Hiermit erkläre ich, dass ich bisher an keiner anderen Hochschule ein Promotionsgesuch eingereicht habe.

Karlsruhe, den 07.01.2014

Acknowledgment

For me, this page is the end of a long journey through hard work, small successes, periods with high and others with few motivation, fruitful discussions and a lot of learning. It was a journey which would not have been possible with the help of the people around me whom I would like to thank at this point.

I am deeply indebted to my advisor Clemens Puppe. Already through my studies and then in my time as his doctoral candidate, he was my constant companion. His great ability to get to the heart of the problems, helped me in numerous discussions. He inspired this thesis with fantastic ideas and helpful comments and enabled me to participate at several international conferences. His positive attitude, his delicate way to announce constructive criticism and his trust in my work pushed me a lot.

I am equally grateful to my second advisor Johannes Philipp Reiß. Although he has spent only a short time at the Department, he supported me enthusiastically with the statistical analysis and helped me with my first steps in STATA. I am very thankful for all the hours of inspiring discussions, his patience and for his helpful and comprehensive explanations.

Furthermore, I would like to thank my external advisor Sebastian Kube for his helpful remarks concerning the experimental section. Special thanks also go to the other members of my PhD Examination Committee, namely Martin Ruckes and Wolf Fichtner.

I greatly benefitted from the fruitful research atmosphere and I would like to thank my highly appreciated room neighbors Caren Söhner and Marten Hillebrand whose doors stood open during all the years for questions and conversations. I have lost count of the number of times, but they never rejected a request when I knocked on their doors to ask for their opinion. A special thank is also due to Tobias Dittrich. With his programming experience he supported this thesis already in his seminar work by implementing our algorithm in JAVA. In addition, I would like to thank Philipp Enders, Claudia Niemeyer and Markus Fels for their helpful remarks. I really enjoyed the atmosphere at the Department and therefore I would like to thank Susanne Fuchs-Seliger, Monika Emanuel, Wolfgang Eichhorn and Andranik Tangian.

The experimental part of this thesis could not have been achieved without Christoph Neumann, Jonas Kohler, Stephan Weiß, Marc Adam, the team in the laboratory, the participation of numerous students and the financial support of the "Fondation Université de Strasbourg".

Furthermore, I would like to thank my former colleagues Tobias Lindner, Johannes Ney, Stefan Rummel and Martin Barbie who supported me in the initial phase of my work. I obtained important input in the field of experimental economics by the participants of the HeiKaMaX and therefore especially like to thank the organizers Joerg Oechssler and Dirk Engelmann. Special thanks go to Christoph Schumacher who helped me with my very first presentation and to Franz Dietrich who helped me to restructure and state more precisely the second chapter of this thesis. In different conferences and discussions, I appreciated the inspiring remarks of Berno Büchel, Klaus Nehring, Salvador Barberà, Hervé Moulin, Berthold Daubner, Ingrid Ott, Patrick Llerena, Herrade Igersheim, Karl-Martin Erhardt, Christian Feige, Jean-François Laslier and many other workshop participants who I have not specifically mentioned.

In addition, I would like to thank my friends who always lent me a sympathetic ear and my volleyball team for the sporting balance. Last but not least, I would like to thank my parents for their never ending love, encouragement and trust. My deepest thank goes to my boyfriend Fabian for his loving support during the last years.

Abstract (English Version)

The topic of my dissertation are generalized single-peaked preferences. The three chapters focus on theory, empirics and experiments, respectively.

In the first chapter, generalized single-peaked preferences are modeled in a space of connected coalitions. We prove a novel possibility result for strategy-proof aggregation of generalized single-peaked preferences using the special structure of the underlying betweenness relation.

In the second chapter, we elicit such preferences empirically by means of a survey. We determine the ordering of political parties such that the largest share of reported preferences are single-peaked over single parties. As a next step, we design and implement an algorithm to check for which ordering of parties the reported preference rankings satisfy *generalized single-peakedness* over connected coalitions. We adapt the model to coalitions which consist of at most two coalition partners. We show that the left-right-spectrum which fits best to our three analyses is different from the common one which is applied as the seating arrangement in the German Parliament.

The third chapter of this thesis reports the results of a laboratory experiment. Using appropriate payoff functions, we induce single-peaked preferences for the participants. Two different voting rules, the mean and the median rule, are investigated and checked for manipulability. We derive equilibria theoretically and compare our experimental data to the equilibrium predictions. We observe that participants play Nash equilibria, but – in contrast to what one might expect – the observed equilibria frequently involve weakly dominated strategies. By introducing manipulation costs, such equilibria are ruled out theoretically and indeed observed less often.

Abstract (German Version)

Gegenstand dieser Dissertation sind verallgemeinert eingipflige Präferenzen. Die Arbeit gliedert sich in drei Kapitel mit jeweiligen Schwerpunkten auf theoretischer, empirischer und experimenteller Vertiefung.

Im ersten Teil werden verallgemeinert eingipflige Präferenzen im Raum der zusammenhängenden Koalitionen modelliert. Es wird ein neues Möglichkeitsresultat zur nicht-manipulierbaren Aggregation von verallgemeinert eingipfligen Präferenzen bewiesen, das die spezielle Struktur der zugrunde liegenden dreistelligen Zwischenrelation ausnutzt.

Im zweiten Teil der Arbeit werden solche Präferenzen empirisch untersucht. Hierfür wurden Umfragedaten erhoben und analysiert. Eine Anordnung der politischen Parteien wird angegeben, so dass der größtmögliche Teil der Befragten eingipflige Präferenzen über einzelnen Parteien hat. Daraufhin wird algorithmisch untersucht, für welche Ordnung die Präferenzen über Koalitionen die *verallgemeinerte* Eingipfligkeit erfüllen. Eine Anpassung des Modells für Koalitionen mit maximal zwei Koalitionspartnern wird vorgenommen. Es zeigt sich, dass sich das Links-Rechts-Spektrum, das in diesen drei Analysen am besten abschneidet, von dem unterscheidet, welches üblicherweise im Bundestag angenommen wird.

Für den dritten Teil der Arbeit wurde ein Experiment durchgeführt, in dem den Teilnehmern mithilfe entsprechend gewählter Auszahlungsfunktionen eingipflige Präferenzen induziert wurden. Zwei Abstimmungsregeln, die auf arithmetischem Mittel bzw. Median basieren, werden auf Manipulierbarkeit untersucht. Neben der theoretischen Herleitung der Gleichgewichte werden die Experimentdaten ausführlich analysiert. Es stellt sich heraus, dass die Teilnehmer zwar Nash-Gleichgewichte spielen, diese jedoch – entgegen der theoretischen Vorhersage – auch schwach dominierte Strategien enthalten. Durch die Einführung geringer Manipulationskosten werden solche Gleichgewichte allerdings ausgeschlossen und tatsächlich auch seltener beobachtet.

Contents

List of Figures

List of Tables

1 Introduction

Several times a day, every one of us has to make different decisions. Some of them are quick and easy as they are taken by one person only and in addition, do not influence any other person. Other decisions are more complex as they may affect the life of another individual and yet others have to be taken by a group of decision makers. In all of the mentioned cases, the preferences of one or more persons are the basis for these decisions.

In the present thesis, we will focus on different aspects of preferences and the way small groups or committees aggregate them. A basic property of the preferences – and therefore part of the title of this thesis – is the assumption of single-peakedness. A collection of preferences is single-peaked if each individual has a favorite alternative such that moving away from that alternative always reduces utility (in an ordinal sense). This definition presupposes a common scale that determines what it means to “move away from” the peak alternative.

The type of this scale depends on the context, as we see by means of the following motivating examples. Imagine that a person buys fresh fruits at the marketplace and needs one kilo of apples to prepare his favorite apple pie. In this situation, a bag with 990 grams is preferred to a bag with 940 grams. Here, the scale of alternatives is given by the natural ordering of real numbers and the actual choice can be easily deduced from the preferences. Such decisions are not the focus of this thesis.

The underlying scale can also be derived from a political spectrum, for instance by sorting parties from liberal to conservative. In many parliamentary elections, however, two or more parties are necessary to form the government. Hence, we sometimes assume that the voters have preferences not only over single parties but also over political coalitions. In the first chapter of this thesis, we consider the case of preferences over *connected coalitions*, i.e. coalitions of parties which form an interval on the given political spectrum. These preferences are no longer single-peaked in the classical sense and we therefore apply the more general concept of *generalized single-peaked* preferences. We prove a novel possibility result for strategy-proof aggregation of these preferences in the first part of the thesis, which is an extended version of the published article “Efficient and Strategy-Proof Voting over Connected Coalitions: A Possibility Result” (Block, 2010).

In these examples, the ordering (“scale”) of alternatives is exogenously given. However, in other cases it is not always clear which ordering to apply, as one may notice when trying to sort the political parties in Germany. If the Social Democrats are to the left of the Union party, how do we arrange the Green party? In the second chapter, we address this question by means of a survey and elicit preferences over political coalitions. We determine the ordering of political parties such that the largest share of reported preferences is single-peaked over *single* parties. As a next step, we design and implement an algorithm to check for which ordering of parties the reported preference rankings satisfy generalized single-peakedness over connected coalitions. It turns out that the ordering obtained is different from the common one which is applied in the seating arrangement in the German Parliament. Furthermore, we adapt our model to coalitions which consist of at most two coalition partners, as this might be plausible for the formation of a government.

In our last example, a family with several members who are spread across the country plans to spend the Christmas holidays together. None of them likes long journeys and therefore each family member prefers to celebrate close to his own home. Here, the alternatives are ordered geographically. This example is more challenging than the first one, as an interaction of the family members is necessary to come to a decision. How they could solve this problem by an appropriate voting mechanism is discussed in the third chapter of this thesis. There, we report the results of a laboratory experiment where two different voting rules, the mean and the median rule, are investigated and checked for manipulability. Already from a theoretical point of view, the two rules are fascinating as the structure of their equilibria is distinct. The mean rule has a unique Nash equilibrium whereas there exist different types of Nash equilibria under the median rule. Although truth-telling is the strategy one might expect, we observe equilibria which involve weakly dominated strategies. We introduce manipulation costs to rule out such equilibria theoretically. Whether these costs influence the behavior of the participants in the laboratory, is discussed in the third chapter of this thesis.

Each chapter of this thesis has a different focus and can be read independently. To facilitate readability, the conclusions and introductions are connecting passages between the chapters.

Part I

Efficient and Strategy-Proof Voting over Connected Coalitions: A Possibility Result

2 Motivation

By the Gibbard-Satterthwaite-Theorem (see Gibbard (1973) and Satterthwaite (1975)) the only strategy-proof voting rule on an unrestricted preference domain over at least three alternatives is the dictatorship of one individual. For possibility results restrictions of the preference domain are necessary. Well-known examples are the domain of all single-peaked preferences on a line (see Moulin (1980)) and the domain of all separable preferences on the hypercube (see Barbera et al. (1991)).

In this chapter a novel example of a possibility domain is presented. As the two preference domains mentioned above (and a number of other possibility domains as well) it belongs to the large class of generalized single-peaked domains considered in Nehring and Puppe (2007b).

To motivate our preference domain, consider a finite set of political parties ordered from left to right on the political spectrum. The space of alternatives is the family of all *connected* coalitions, i.e. the family of all non-empty coalitions that contain with any two parties all parties that are between them in the political spectrum. The family of connected coalitions can be endowed with a natural betweenness relation as follows: a connected coalition C is *between* two connected coalitions C_1 and C_2 if (i) the leftmost element of C is between the leftmost elements of C_1 and C_2, respectively, and (ii) the rightmost element of C is between the rightmost elements of C_1 and C_2, respectively.

A preference ranking on the family of all non-empty connected coalitions is called *generalized single-peaked* if it admits a unique most preferred coalition (the "peak"), say C^*, such that a coalition C is strictly preferred to another coalition C' whenever C lies between C' and C^*.

We show that on the domain of all generalized single-peaked preferences over connected coalitions there exist anonymous and strategy-proof social choice functions. One example is the social choice function that selects the connected coalition which has as leftmost element the median of the leftmost elements of the individually most preferred coalitions and as rightmost element the median of the rightmost elements of the individually most preferred coalitions.

The existence of anonymous and strategy-proof social choice functions on the domain of all generalized single-peaked preferences over connected coalitions follows from general results derived in Nehring and Puppe (2007b) since the underlying betweenness relation gives rise to a *median space*. In fact it follows from the analysis in Nehring and Puppe (2007b) that the social choice function described above is the *only* anonymous and strategy-proof voting rule that is neutral in an appropriate sense. Moreover, using the main result of Nehring and Puppe (2007a) one can show that the above voting rule is *efficient*. In this chapter, we provide elementary proofs of its strategy-proofness and efficiency.

This chapter is the extended version of the published article “Efficient and Strategy-Proof Voting over Connected Coalitions: A Possibility Result” (Block, 2010).

3 Generalized Single-Peaked Preferences over Connected Coalitions

Let $A = \{p_1, \ldots, p_m\}$ be a finite set containing $m \geq 2$ objects. We consider the case in which individuals have preferences over a subset of the power set $\mathcal{P}(A)$. Specifically, we consider the following domain restriction. Let $<$ be a linear ordering of A, w.l.o.g. $p_1 < \cdots < p_m$. As a specific example one may think of A as representing a set of political parties which can be ordered from left to right on a political spectrum. In this case the power set $\mathcal{P}(A)$ represents the class of possible coalitions. While other interpretations may be applicable as well, in the remainder will refer to the elements of A as political parties and to the elements of $\mathcal{P}(A)$ as coalitions. For notational convenience we identify parties with their indices and simply write (ijk) for $\{p_i, p_j, p_k\}$. A non-empty coalition C is called **connected**, if for all i, j, k,

$$i, j \in C \text{ and } i < k < j \Rightarrow k \in C.$$

We denote by $\mathcal{C}_< \subset \mathcal{P}(A)$ the set of all connected coalitions.

For every connected coalition C we call

$$l_C \in C \text{ **leftmost in C** if for all } k < l_C \Rightarrow k \notin C$$

and

$$r_C \in C \text{ **rightmost in C** if for all } k > r_C \Rightarrow k \notin C.$$

Evidently one has $C = (l_C ... r_C)$ for every connected coalition C.

Example 1. The coalition $C = (234)$ consisting of the parties p_2, p_3 and p_4 is connected with $l_C = 2$ and $r_C = 4$. $C' = (24)$ is not connected and therefore not an element of $\mathcal{C}_<$.

We define the following betweenness relation on $\mathcal{C}_<$. A coalition C is **between** C_1 and C_2 if

- its leftmost party l_C is between l_{C_1} and l_{C_2}, and
- its rightmost party r_C is between r_{C_1} and r_{C_2}.

Formally, C is between C_1 and C_2 if

$$l_C \in [\min\{l_{C_1}, l_{C_2}\}, \max\{l_{C_1}, l_{C_2}\}] \text{ and } r_C \in [\min\{r_{C_1}, r_{C_2}\}, \max\{r_{C_1}, r_{C_2}\}].$$

For graphical illustration of the betweenness relation consider Figure 3.1 with $m = 6$ parties. A coalition C is between C_1 and C_2 if and only if it lies on a shortest path connecting C_1 and C_2 on the graph. Note that shortest paths need not be unique.

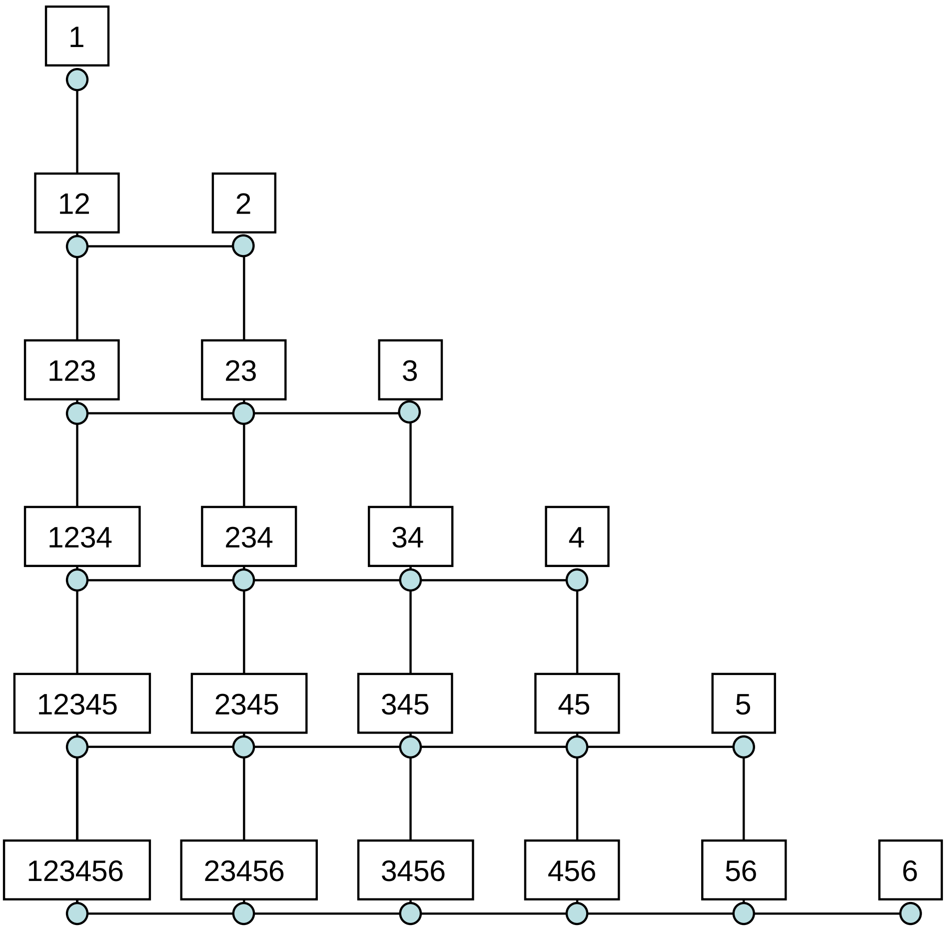


Figure 3.1: Graphical illustration for six parties

Obviously, the betweenness relation respects the subset ordering, i.e. a coalition is between any of its subsets and any of its supersets (see Figure 3.2). The neighbors of a coalition $(l_C \ldots r_C)$ are all coalitions which are connected and which consist either of exactly on party more $((l_C - 1 \ldots r_C)$ or $(l_C \ldots r_C + 1))$ or one party less $((l_C \ldots r_C - 1)$ or $(l_C + 1 \ldots r_C))$.

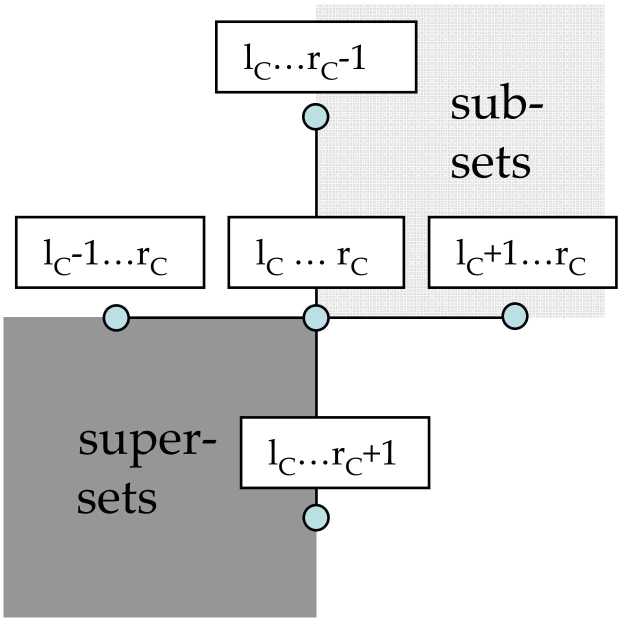


Figure 3.2: The coalition C with its neighbors

We are now able to define the preference structure over connected coalitions. Let $N = \{1, \ldots, n\}$ denote the set of voters. Suppose that every individual i has a unique favorite coalition $C_i^* = (l_{C_i}^* \ldots r_{C_i}^*)$ which is called **peak of i**, i.e. for all $C \in \mathcal{C}_<$

$$C \neq C_i^* \Rightarrow C_i^* \succ_i C.$$

The preference relation ($\succsim_i$) of individual i is **generalized single-peaked** if for all connected coalitions C and $C' \neq C_i^*$ we have that

$$C \text{ is between } C_i^* \text{ and } C' \Rightarrow C \succ_i C'.$$

Denote by $\mathcal{S}(\mathcal{C}_<)$ the set of all generalized single-peaked preferences on $\mathcal{C}_<$.

Example 2. Suppose that $\succ$ is generalized single-peaked with peak $C_i^* = (234)$. Then, for instance, $(234) \succ_i (23) \succ_i (123)$, but there is no restriction on the preference over (23) and (1234).

Remark: The concept of generalized single-peakedness over connected coalitions cannot be reduced to single-peakedness in the classical sense, i.e. there does not exist a linear ordering of the set of connected coalition such that all elements of $\mathcal{S}(\mathcal{C}_<)$ are single-peaked with respect to the given linear ordering in the classical sense.[1]

A social choice function is a mapping

$$F := \begin{cases} \mathcal{S}(\mathcal{C}_<)^n & \longrightarrow \quad \mathcal{C}_< \\ (\succsim_1, \dots, \succsim_n) & \longmapsto \quad C \end{cases}$$

F is called **strategy-proof** if for all $i \in N$ and $\succsim_i, \succsim_i' \in \mathcal{S}(\mathcal{C}_<)$:

$$F(\succsim_1, \dots, \succsim_i, \dots, \succsim_n) \succsim_i F(\succsim_1, \dots, \succsim_i', \dots, \succsim_n)$$

[1] To see this, consider three generalized single-peaked preferences $\succsim_i, \succsim_j, \succsim_k$ which have the same peak but pairwise different second-best coalitions. Evidently, it is not possible to arrange all three second-best coalitions as direct neighbors of the peak in one dimension.

4 Results

Consider the social choice function

$$F(\succsim_1, \ldots, \succsim_n) = (med(l_{C_1^*}, \ldots, l_{C_n^*}) \ldots med(r_{C_1^*}, \ldots, r_{C_n^*})),$$

where med denotes the median-operator, i.e. $med(x_1, \ldots, x_n)$ is an element satisfying

$$\#\{i \mid x_i \leq med(x_1, \ldots, x_n)\} \geq \frac{n}{2} \text{ and } \#\{i \mid x_i \geq med(x_1, \ldots, x_n)\} \geq \frac{n}{2}.$$

For simplicity we assume here that the number of voters is odd. This guarantees that the median-operator produces a single element (for further discussion see the remark at the end of this section).

Theorem 1.
$F(\cdot)$ is strategy-proof and anonymous.

Proof:
F is anonymous:
The median operator is anonymous, hence F is anonymous as well.

F is strategy-proof:
Let $F(\cdot) = C^* = (l^* \ldots r^*)$ be the social choice.
Suppose that F is not strategy-proof. Then, there exists a misrepresentation $\succsim_i'$ in $\mathcal{S}(C_<)$ such that $F(\succsim_1, \ldots, \succsim_i', \ldots, \succsim_n) \succ_i C^*$ for an individual i with peak $C_i^* = (l_{C_i}^* \ldots r_{C_i}^*)$. As F depends only on the peak profile it follows that $F(\succsim_1, \ldots, \succsim_i', \ldots, \succsim_n) = C^*$ for all $\succsim_i'$ with peak C_i^*. So the peak of $\succsim_i'$ has to be different from C_i^*. Let $C' = (l' \ldots r')$ be this misrepresented peak and $F(\succsim_1, \cdots \succsim_i', \cdots \succsim_n) = C_M' = (l_M' \ldots r_M')$ the resulting (manipulated) social choice.

For the relative position of the leftmost elements $l^*, l_{C_i}^*, l'$ there are the following three possible cases:

Case 1) l' is between l^* and $l_{C_i}^*$ $\Rightarrow$ the median of the leftmost elements does not change.

Case 2) $l_{C_i}^*$ is between l' and l^* $\Rightarrow$ the median of the leftmost elements does not change.

Case 3) l^* is between $l^*_{C_i}$ and $l' \Rightarrow$ the median of the leftmost elements is between l^* and l'. This implies that l^* is between $l^*_{C_i}$ and l'_M.

Notice that since betweenness is always understood in the weak sense we have that l^* is between $l^*_{C_i}$ and l'_M also if the median of the leftmost elements does not change, i.e. if $l^* = l'_M$.

Analogously, one easily shows that r^* is between $r^*_{C_i}$ and r'_M. This implies that C^* is between C^*_i and C'_M, hence by generalized single-peakedness, $C^* \succsim_i C'_M$. Thus, i has no incentive to misrepresent. □

Example 3. There are $m = 6$ parties and $n = 5$ individuals with peaks on the coalitions $C^*_1 = (2), C^*_2 = (123), C^*_3 = (34), C^*_4 = (45), C^*_5 = (23456)$ respectively. The median of the leftmost parties is $med\{2,1,3,4,2\} = 2$ and the median of the rightmost parties is $med\{2,3,4,5,6\} = 4$. Therefore, $F(\succsim_1, \dots, \succsim_n) = (234)$.

Remark: It follows from the analysis of Nehring and Puppe (2007b) that the social choice function $F(\cdot)$ given above is the only anonymous and strategy-proof voting rule on $\mathcal{S}(\mathcal{C}_<)^n$ that is *neutral*, i.e. that treats the elements of $\mathcal{C}_<$ symmetrically (in an appropriate sense). As shown in Nehring and Puppe (2007a) the property of neutrality is closely related to efficiency, to which we turn now.

Proposition 2.
$F(\cdot)$ is efficient, i.e. for all $(\succsim_1, \dots, \succsim_n)$ there exists no $C \in \mathcal{C}_<$ such that $C \succsim_i F(\succsim_1, \dots, \succsim_n)$ with at least one strict preference.

Proof: Consider a situation where $F(\succsim_1, \dots, \succsim_n) = C^* = (l^* \dots r^*)$ and let $O_1 := \{C \in \mathcal{C}_< | l_C \geq l^* \text{ and } r_C \leq r^*\}$, $O_2 := \{C \in \mathcal{C}_< | l_C \leq l^* \text{ and } r_C \leq r^*\}$, $O_3 := \{C \in \mathcal{C}_< | l_C \leq l^* \text{ and } r_C \geq r^*\}$, $O_4 := \{C \in \mathcal{C}_< | l_C \geq l^* \text{ and } r_C \geq r^*\}$ (see Figure 4.1).

Step 1: We prove that there exists a peak in every orthant $O_1, \dots, O_4$ by contradiction. By symmetry, we assume w.l.o.g. that there is no peak in O_1, i.e. for all $i = 1, \dots n$, $l^*_{C_i} < l^*$ or $r^*_{C_i} > r^*$.

As l^* is the median of the leftmost elements, we have $\#\{i | \, l^*_{C_i} < l^*\} < \frac{n}{2}$ (remember that n is odd). By the same argument, $\#\{i | \, r^*_{C_i} > r^*\} < \frac{n}{2}$.

Summing up both inequalities, we obtain:
$\#\{i | \, l^*_{C_i} < l^* \text{ or } r^*_{C_i} > r^*\} < n$, a contradiction; thus O_1 contains a peak.

Step 2: If there is a peak in every orthant, then C^* is efficient.

Case a) Evidently, C^* is efficient if $C^*_i = C^*$ for some i.

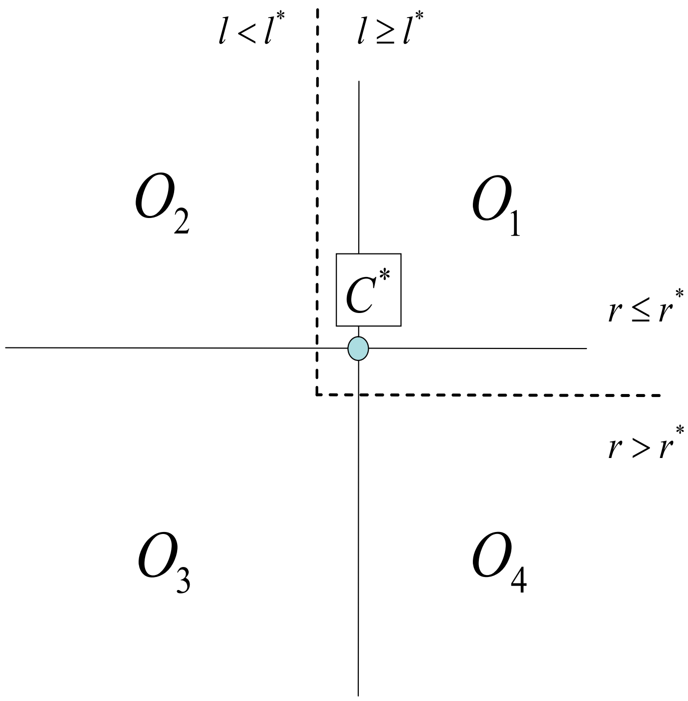


Figure 4.1: The social choice C^* and the orthants $O_1, \dots, O_4$

Case b) No individual has his peak on C^*.
Let $\hat{C} \neq C^*$ be an arbitrary connected coalition, say in orthant O_a. By Step 1 we know that there exists a peak $C_j^* \in O_{a+2(mod4)}$ in the opposite orthant. As C^* is between $\hat{C}$ and C_j^*, it follows by generalized single-peakedness that $C^* \succ_j \hat{C}$. Thus, C^* is efficient. □

Remark: If the number of voters is even, the median operator does not always assign a single element, i.e. the mapping F introduced above is a correspondence. To obtain a strategy-proof social choice *function* in that case one has to give up either anonymity or neutrality. Neutrality can be maintained, for instance, by counting the peak of one pre-specified individual twice; clearly, this violates anonymity. Anonymity, in turn, can be maintained by adding phantom voters along the lines suggested by Moulin (1980). For example, one may add a phantom voter with peak at the grand coalition $C = (1 \dots m)$. The resulting social choice function selects the largest connected coalition among all median elements. Obviously this rule is not neutral since it does not treat alternatives symmetrically.

Part II

Preferences over Political Coalitions: An Empirical Study on Generalized Single-Peakedness

5 Motivation

In this chapter we relate conclusions from Social Choice Theory to results from an empirical survey. In Social Choice Theory the assumption of single-peaked preferences is used often but only little empirical confirmation of this assumption has yet been given.

We consider three different kinds of preferences: Single-peaked preferences over singletons, generalized single-peaked preferences over connected coalitions and generalized single-peaked preferences over "small" coalitions. The underlying domain of alternatives is a linearly ordered set, e.g. a political left to right ordering. In case of generalized single-peaked preferences a multi-dimensional space is needed. Therefore, we consider the structure of median spaces, introduced by Nehring and Puppe (2007b) and applied also by Block (2010).

The analysis of the preferences is based on two empirical studies in which students were asked to rank political coalitions according to their preferences. Both studies were conducted in Karlsruhe, the first in Spring 2009, the second in Winter 2009. In between these two surveys the 17th German Parliament was elected.

The two studies lead to surprisingly similar findings, despite the significant political change brought about by the election.[2] Firstly, preferences over political coalitions remain stable over time and are roughly the same in both surveys. Secondly, the underlying order with respect to which the maximal number of preferences are single-peaked stayed the same, which can be interpreted as saying that people's perception and conceptualization of parties was not perturbed much. This stability increases the trust that one may have in the robustness of the findings.

Concerning the single-peakedness assumption on preferences over political parties, we identify an ordering of political parties such that most of the respondents have single-peaked preferences over parties. This ordering is the same for both groups of the survey. In the second part of this work, we analyze the structure of preferences over coalitions. We consider two special cases: *Connected coalitions* are sets of political parties which form an interval in the left-to-right ordering. *Small coalitions* are coalitions consisting of one or two political parties.

[2]The pre-election parliament gave rise to a grand coalition of the Union and SPD, whereas the post-election parliament led to a centre-right coalition.

For single-peakedness, Escoffier et al. (2008) give an algorithm to determine whether a profile is single-peaked with respect to some dimension. To the best of my knowledge, for generalized single-peakedness, there exists no such algorithm. With the help of a program implemented in JAVA, our data are analyzed. As the sample has a size of approximately 500 respondents, it is likely that there exists no structure such that all preferences fit into the generalized single-peaked pattern. Hence, a preference ranking is called *almost* generalized single-peaked if at most four change iterations are necessary until the preferences are generalized single-peaked. We found an ordering of political parties such that 81% of preferences fit into the model of almost generalized single-peaked preferences over connected coalitions. A similar analysis was done for small coalitions and there exists an ordering of political parties such that about 71% of preferences are almost generalized single-peaked. The ordering of political parties is the same for all three results.

Finally, we discuss the comments given by the students as to the influence of extreme parties and their preferences for coalitions of a particular size.

In this part of my thesis, combinatorial considerations are used as well as (computational) social choice results. All conclusions are based on empirical data and compared with theoretical assumptions.

6 Overview of the literature

In this section, we give a short overview over the existing literature. We found theoretical results analyzing single-peaked preferences and empirical studies describing the structure of preferences over political parties. To the best of my knowledge, there is no empirical analysis over the preference structure over political coalitions yet.

The theory of single-peaked preferences has a long tradition. Black (1958) analyzed this type of preferences for the first time: If voters' preferences are single-peaked with respect to a given order, the median alternative will receive a simple majority against every other alternative and no Condorcet cycle (Condorcet, 1785) will occur. The median rule satisfies the property of strategy-proofness which was shown by Moulin (1980). He also gave a detailed description of all strategy-proof voting rules on single-peaked domains in the one-dimensional case.

However, the probability that a preference profile satisfies the assumption of single-peakedness is decreasing as well in the number of individuals as in the number of alternatives when alternatives are assumed to be uniformly distributed in a preference ranking. This is one of the theoretical results already stated by Niemi (1969). For many years the characterization of preferences has still been an interesting field of research. Ballester and Haeringer (2007) characterize the (one-dimensional) single-peaked domain and prove a possibility result whose reformulation has serious implications: "If a profile is not single-peaked, then there must be a violation of that property for a set of three preferences over three alternatives or a set of two preferences over four alternatives." Therefore, it is almost impossible to find a single-peaked profile in practice when considering large groups.

In "A Theory of Data", Coombs (1960) describes the "Unfolding Theory" to transform scales. In particular, from one individual preference ranking all (common) orderings of alternatives are derived, such that the preference ranking is single-peaked. A measure for "proximity to single-peakedness" is explained by Niemi (1969) and often applied. The share of single-peaked preferences in the entire profile is calculated and maximized over all possible or reasonable orderings. Accordingly, in our study, we try to find an ordering, such that the largest proportion of preference rankings is single-peaked

Single-peakedness is not only a problem in economics but also a challenge in computer science. Escoffier et al. (2008) give an algorithm to determine whether a profile is single-peaked with respect to some axis. They call this property "single-peaked-consistency" and calculate the running time of the algorithm. With the running time $O(nm)$, the algorithm is applicable also with high numbers of voters and is adaptable when new voters are added to the profile. In this context the connection to artificial intelligence problems becomes clear. The probability of single-peaked consistency decreases exponentially both with the number of voters and the number of candidates. As approximation, they suggest different approaches: First, to delete the minimal number of individual rankings, such that single-peakedness is satisfied (similar to the proximity mentioned above) or second, to delete the minimal number of candidates, i.e. alternatives, such that the profile is single-peaked and third, the minimal number of axis, such that each individual preference is single-peaked with respect to at least one axis. Note that the latter is only a list of one-dimensional axis and still not multi-dimensional. Escoffier, Lang and Öztürk calculate the minimum and maximum number of axes that are compatible with a set of distinct votes and candidates. The underlying profile is uniformly distributed. In our work, we restrict preferences to real ones collected in a survey but we compare the results with the uniformly distributed case.

How can the theoretically obtained results be applied in a political context? Niemi and Wright (1987) analyzed nationally-representative samples of preferences over presidential candidates of the U.S. in 1980. With 14 politicians and computers from those times, they had to divide their sample in smaller sub-sets to determine the proportion of single-peaked preferences. However, they found that "[the] extreme popularity or unpopularity of candidates leads to a high degree of unidimensionality, but the underlying dimensionality is not ideologically based." Radcliff (1993) used data of U.S. elections, too. In contrast to Niemi and Wright (1987), he only considered the "conventional left-right ideological dimension", i.e. he considered an a priori defined ordering. With five candidates, 50% of the preferences were single-peaked.

Van Deemen and Vergunst (1998) ran a national parliamentary election study during Dutch elections in 1994. About 1500 persons were asked to evaluate political parties by assigning (agreement) points. From these agreement points, individual preference rankings were deduced. The main result of their study was the non-existence of Condorcet cycles in this context. Furthermore, they give a linear ordering of eight parties such that 35.1% of the voters have single-peaked preferences. This percentage was maximal compared to all other orderings. Van Deemen and Vergunst (1998) conclude that "single-peakedness of voters preferences in Dutch elections is not a very likely cause" for the observed stability.

Whether preferences and the underlying ordering of alternatives are stable is one focus of the research of List et al. (2013) and Farrar et al. (2010). The measure for proximity to single-peakedness is applied to determine the effects of deliberation on a common ordering. Their main results are similar: "Deliberation can prevent majority cycles [...] by bringing preferences closer to single-peakedness" (List et al.,

2013) “Deliberation effects policy attitudes and brings policy preferences closer to single-peaked preferences.” (Farrar et al., 2010).

We will not discuss here, what influence deliberation may have had on our data. However, it is an interesting and even philosophical question whether there exists an underlying objective ordering of the alternatives which has to be discovered in a process of deliberation. Examples for a left-right dimension in which party competition takes places is the degree of government intervention in the economy (Downs, 1957). Some consequences are that “[p]arties that are ambiguous in their stance on the Left-Right scale attract lower preferential evaluations from individual respondents and, [...], individuals’ level of misrepresentation of a party’s left-right position is inversely related to their preference for that party” according to a study by Aldrich et al. (2010).

So far, the mentioned literature has focused on single parties on the line only. Coalitions are often seen from a game-theoretical perspective, i.e. the candidates have to make the decisions. For instance, Brams et al. (2002) consider a coalition formation process for connected coalitions. Each voter is part of the coalition forming alternatives. In our model, however, we consider individuals who have preferences over distinct political parties.

One empirical study on whom voters vote for if they have preferences over coalitions but only one vote for a party was run by Blais et al. (2006). They compare different elections were political parties announced possible coalition partners before the election. The behavior of voters was influenced by these promises when they decided to vote for one party. “For one out of ten, coalition preferences were a decisive consideration, i.e. they induced the voter to support a party other than their most preferred one.”

A multi-dimensional approach was formulated by Nehring and Puppe (2007b) in which preferences are no longer single-peaked with respect to one axis. In the previous part of this thesis and in Block (2010) a model is developed which embeds connected coalitions in the median space structure. We will discuss the application of this model in detail for the concrete application of the parties in the German Parliament.

7 Theoretical model

In this section, we focus on the theoretical model. Hence, we provide a structure of coalitions on a graph and define preference relations on this graph. These definitions are abstract at first, but we apply them to concrete examples in the next section.

We consider a finite set of parties $A = \{p_1, ..., p_m\}$, where m is the number of parties. In our empirical application, A will contain the $m = 5$ parties represented in the 17th German Federal Parliament, namely the Union (U)[3], the Social Democratic Party (S), the Free Democratic Party (F), the Left (L) and Alliance'90/The Greens (G).[4]

The power set $\mathcal{P}(A)$ without the empty-set is called the **set of coalitions**. Let $\mathcal{C}_M \subset \mathcal{P}(A)$ be a subset of $\mathcal{P}(A)$ which satisfies property $M \in \{single, con, small\}$. We consider sets of coalitions with three different properties, which are explained in detail in the corresponding subsections:

$$
\begin{aligned}
\mathcal{C}_{single} &= \text{the set of coalitions which consist of a single party} \\
\mathcal{C}_{con} &= \text{the set of connected coalitions} \\
\mathcal{C}_{small} &= \text{the set of small coalitions}
\end{aligned}
$$

Let $\mathcal{C}_M = (\mathcal{C}_M, \mathbb{E}_M)$ be a connected graph[5] consisting of $|\mathcal{C}_M|$ vertices and a set of edges $\mathbb{E}_M$. An edge is a tuple $\{C_1, C_2\}$ which *links* two different coalitions $C_1, C_2 \in \mathcal{C}_M$.

A path from a coalition C in $\mathcal{C}$ to another D is a finite sequence of coalitions $(C_1, C_2, ..., C_k)$ in $\mathcal{C}$ such that $C_1 = C$, $C_k = D$, and any adjacent coalitions are linked, i.e., $\mathbb{E}$ contains each pair $\{C_t, C_{t+1}\}$. A shortest path from C to D is a path from C to D of minimal length. A coalition C is between coalitions C_1 and C_2 if $\mathcal{C}$ belongs to a shortest path from C_1 to C_2.

[3]To be legally precise, the Union is not a party but a parliamentary faction consisting of two sister parties, the Christian Democratic Union (which covers the non-Bavarian part of Germany) and the Christian Social Union (which covers Bavaria).

[4]For more information and historical remarks about the German Federal Parliament see Appendix A.

[5]A graph is called connected if for every pair of vertices $C, D \in \mathcal{C}_M$, there exist a path from C to D.

Note: Up to now we have made no restrictions as to which edges belong to $\mathcal{C}_M$. An edge might represent similarities. In the following subsections we give three examples.

Therefore, we fix a property $M \in \{single, con, small\}$ and consider the preferences over the coalitions $\mathcal{C}_M$.

We consider a group of individuals, labeled $i = 1, 2, \ldots, n$, where n is the (finite) group size. Each individual has a preference ranking $(\succsim_i)$ over the set of coalitions $\mathcal{C}$, which is assumed to be a linear order.[6] The corresponding relation of strict preference is denoted as $(\succ_i)$; formally, $C \succ_i C'$ if and only if $C \succsim_i C'$ and $C \neq C'$. The **peak** of i's preference ranking $(\succsim_i)$, or in short of individual i, is the (unique) most preferred coalition, i.e. the $C_i^* \in \mathcal{C}$ such that $C_i^* \succsim_i C$ for all $C \in \mathcal{C}$. The preference ranking is **generalized single-peaked** with respect to $\mathcal{C}_M$, if $C' \succsim_i C$ for each coalition C and each coalition C' between C and i's peak. Accordingly, we call a profile $(\succsim_1, \ldots, \succsim_n)$ generalized single-peaked with respect to $\mathcal{C}_M$, if each of the n preference relations is generalized single-peaked with respect to the same $\mathcal{C}_M$.

In the following, we denote coalitions simply by listing its members: the coalition $\{1, 2\}$ is abbreviated 12, the singleton coalitions $\{3\}$ is abbreviated 3, and so on.

7.1 Single-peaked preferences over single-party coalitions

We consider preferences over the set of single-party coalitions $\mathcal{C}_{single} = \{P : P \in A\}$. We identify each coalition P in $\mathcal{C}_{single}$ with the party P, so that the domain of preferences becomes simply the set of parties $\mathcal{C}_{single} \equiv A$. In order to represent the standard notion of single-peakedness graph-theoretically, we endow $\mathcal{C}_{single}$ with a linear graph structure, as illustrated in Figure 7.1. Formally, $\mathbb{E}$ consists of all pairs $\{C, D\}$ of adjacent parties with respect to some fixed linear order of parties. Therefore, $|\mathbb{E}| = m - 1$. Generalized single-peaked preferences on $\mathcal{C}_{single}$ reduce to the so-called *single-peaked preferences* which are well-known in decision theory.

So far, we have only fixed the cardinality of $\mathbb{E}$, but not determined the edges contained in $\mathbb{E}$, i.e. which parties are adjacent or in other words which is the underlying linear order of the political parties. There are many interpretations of the resulting

[6] In this thesis, different "things" have to be ordered. To prevent confusion, we will use the term "(preference) ranking" for preferences and the term "order" or "ordering" for political parties.

left to right ordering which depends on dimensions like political concern, foreign policy, diversity, freedom, change or social power.[7]

In the following, we consider all possibilities of linear orderings and try to find which of them is adequate for the example of German political parties. From the combinatorial point of view, there are $m!$ possibilities to order m parties on the line. As the resulting graph is symmetrical, this number can be reduced to $\frac{m!}{2}$. In our example of $m = 5$ parties, we get 60 different orderings. Two of them are shown in Figure 7.1 and 7.2. The only preference rankings that are single-peaked with respect to both orders are $2 > 1 > 3 > 4 > 5$ and $1 > 2 > 3 > 4 > 5$. For more combinatorial remarks, see Appendix C.

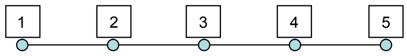


Figure 7.1: Ordering 12345

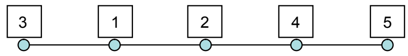


Figure 7.2: Ordering 31245

7.2 Generalized single-peakedness

We now turn to preferences over a different set of coalitions than the set of single-party coalitions. In Section 7.2.1 we consider the set of connected coalitions, and in Section 7.2.2 the set of "small" coalitions. We do not assume such preferences to satisfy standard single-peakedness w.r.t. a linear order of the coalitions; such a linear order would not have been very plausible, since a coalition can include parties from the entire political spectrum. Rather, we assume preferences to be generalized single-peaked with respect to a graph over coalitions.

7.2.1 Generalized single-peakedness over connected coalitions

In this example, we focus on generalized single-peaked preferences over connected coalitions, as introduced in Block (2010).[8] As in the former section, we assume that a linear order $<$ of the five parties is given, i.e. $p_1 < \ldots < p_5$ with $p_i \in A$.

A non-empty coalition C is called **connected** with respect to this linear order if whenever C contains p_i and p_j, then C also contains p_k with $i < k < j$. We denote

[7] "The terms Right and Left refer to political affiliations which originated early in the French Revolutionary era of 1789-1796, and referred originally to the seating arrangements in the various legislative bodies of France. The aristocracy sat on the right of the Speaker (traditionally the seat of honor) and the commoners sat on the Left, hence the terms Right-wing politics and Left-wing politics." Source: `http://en.wikipedia.org/wiki/Political_spectrum`, retrieved December 2013.

[8] Actually, the graph with its two dimensions can be interpreted as a property space, see Nehring and Puppe (2007b).

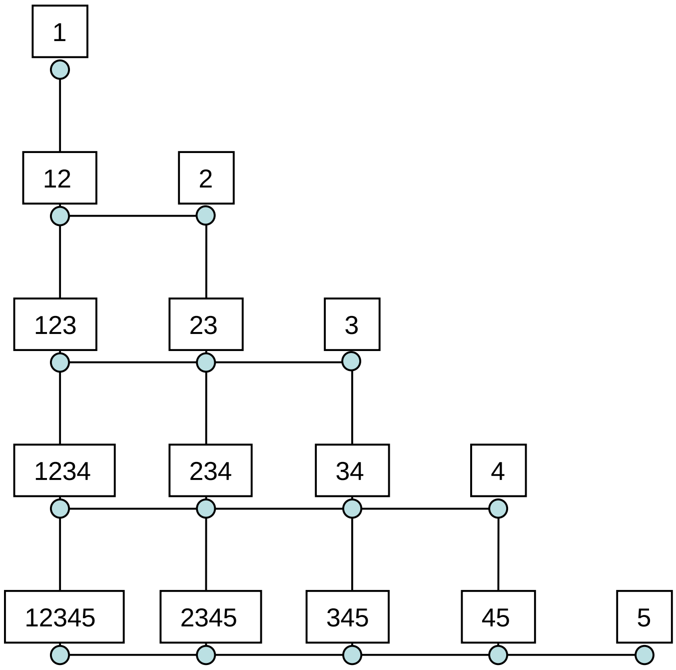


Figure 7.3: Embedding connected coalitions in a median space structure

the set of connected coalitions by $\mathcal{C}_{con}$. In the case of $m = 5$, there are exactly 15 coalitions in $\mathcal{C}_{con}$. For every connected coalition C we call $l_C \in C$ leftmost in C if for all $p_k < l_C$ we have $p_k \notin C$ and $r_C \in C$ rightmost in C if for all $p_k > r_C$ we have $p_k \notin C$. In case of C being a singleton $l_C = r_C$. For $\mathcal{C}_{con}$, the betweenness-relation, which we have defined in the previous paragraph, reads as follows: A coalition C is between C_1 and C_2, if its leftmost party is between l_{C_1} and l_{C_2} and its rightmost party is between r_{C_1} and r_{C_2}.[9]

The idea of connected coalitions is that two extreme parties cannot form a coalition without accepting moderate parties to be part of the coalition. A justification for this assumption is that voters prefer parties to ally with parties who have similar programs until their coalition is big enough to form a government. Therefore, two connected coalitions are linked by an edge if and only if one of them arises from the other one by either adding or removing the rightmost or the leftmost party.

Example 4. Figure 7.3 illustrates $\mathcal{C}_{con}$ for the order 12345. For simplicity, parties are identified with their indices. For instance, the coalition 12 is connected with respect to the order 12345, but 13 is not (and therefore not shown in the graph).

[9]Formally, C is between C_1 and C_2 if

$$l_C \in [\min\{l_{C_1}, l_{C_2}\}, \max\{l_{C_1}, l_{C_2}\}] \text{ and } r_C \in [\min\{r_{C_1}, r_{C_2}\}, \max\{r_{C_1}, r_{C_2}\}].$$

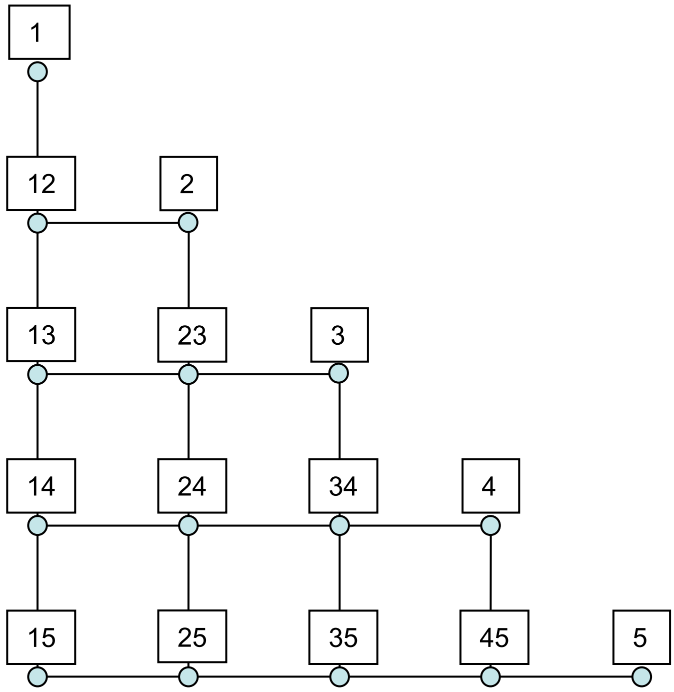


Figure 7.4: Embedding small coalitions in a median space structure

Both 12 and 23 are between the coalitions 123 and 2. We observe, that in the first column the leftmost element of each coalition is party 1 and the number of elements of the coalition increases from the top to the bottom within the row by adding a new rightmost element.

As their are $m!$ different orders, there exist $\frac{m!}{2} = 60$ different graph structures for symmetry reasons, of which only one example is illustrated here.

7.2.2 Generalized single-peakedness over small coalitions

A coalition is *small* if and only if it contains only one or two parties. Let $\mathcal{C}_{small}$ be the set of small coalitions. As in the previous subsection, we consider a linear order $<$ of the parties. Two coalitions $C = (l_C \; r_C)$ and $D = (l_D \; r_D)$ are linked by an edge if and only if one of the following conditions hold.

1. $r_C + 1 = r_D$ and $l_C = l_D$
2. $r_C - 1 = r_D$ and $l_C = l_D$
3. $r_C = r_D$ and $l_C + 1 = l_D$
4. $r_C = r_D$ and $l_C - 1 = l_D$

Figure 7.4 shows the graph for small coalitions for the order 12345. We see, that $|\mathcal{C}_M| = \sum_{i=1}^{m} i$. The graphs $\mathcal{C}_{con}$ and $\mathcal{C}_{small}$ are isomorphic, that is, there exists a bijection $f : \mathcal{C}_{con} \to \mathcal{C}_{small}$ such that two coalitions C and $D \in \mathcal{C}_{con}$ are linked if and only if the corresponding coalitions $f(C)$ and $f(D) \in \mathcal{C}_{small}$ are linked.[10]

[10] Let $f(l_C(l_C+1)\dots(r_C-1)r_c) = (l_C r_C)$, i.e. delete the parties lying in the interior of a coalition, to get an adequate isomorphism.

8 Data collection

The results presented in this chapter are based on two surveys carried out at the Karlsruhe Institute of Technology, Germany (KIT). The first survey (S_1) was conducted in January and February in 2009, the second one (S_2) in November and December in 2009. In between these two surveys, the 17th German Parliament was elected. The online questionnaire for both surveys was identical. It was generated on www.onlineumfragen.com (see Appendix D). A link to the questionnaire was sent to more than one thousand people via e-mail, mostly students at the Department of Economics and Business Engineering. Passwords were used to ensure that everybody filled in only one form. There was no time limit for filling-in. Because of the use of passwords, delays of several days were possible. The average time for answering the question was 509 seconds (8.5 minutes).

In the questionnaire the respondents were asked to rank 31 coalitions according to their political preferences. These 31 coalitions were the non-empty subsets of the power set of the political parties which are represented in the German Parliament (Deutscher Bundestag) at that time. No additional information about the political program of the parties was given. Each respondent had to make a drag and drop list beginning with his favorite coalition and ending with his least desired one. In addition, a text field offered the possibility to give comments. We do not take into account the relative weight (fraction of votes) of parties.

In the first survey 306 people participated of which 266 completely filled-in the questionnaires.[11] In the second survey, there were 278 and 250 participants, respectively. In the remaining analysis, only the completely filled-in questionnaires are considered. We refer to one dataset as a list, which we also call a *ranking*, of 31 elements ordered from the most preferred to the least preferred political coalition.

For the analysis, data were anonymized and comments were detached. For easier handling, coalitions were renamed to numbers from 1 to 31.

[11] Complete data and a translation of the questions are available from the author on request.

9 Method

The aim of the following analysis is to check whether there exists an ordering such that a profile is generalized single-peaked. We use the data obtained by the survey as profile. As explained in Section 7 the property of generalized single-peakedness depends on the underlying graph $\mathcal{G}_{\mathcal{C}_M}$. For each of the three properties $M \in \{single, con, small\}$ and each underlying order of parties, we calculate the percentage of voters that have generalized single-peaked preferences. The methods to obtain these percentages differ for property *single* and (*con* and *small*).

9.1 Method: Single-peaked preferences over singletons

This paragraph describes a method to check which preferences are single-peaked with respect to one fixed order of political parties. For solving this problem, we use the "unfolding technique" introduced by Coombs (1960). The main idea of this technique is as follows: The party which is ranked worst has to be at one extreme of the axis. The penultimate party therefore has to be at the other extreme or a direct neighbor of the last party. Applying recursively this algorithm we get 2^{m-1} preference rankings that are single-peaked with respect to the fixed linear order.[12] In our case, for each order of the five political parties, there are 16 preference relations that are single-peaked, which we call the **reference set** (see Appendix C).

To compare empirical data with the reference set, we first reduce preference rankings over coalitions to preference rankings over parties, e.g. lists of five elements. Then, we determine the percentage of preference rankings which are elements of the reference set. As a last step, we sort the orderings starting with the one that has achieved the highest percentage to find the ordering for which most of the preferences are single-peaked. For our specific data, results are given in Section 10.

In theory, we may ask the question, whether there exist an ordering such that the complete profile, i.e. each of the n individual preference relations, is single-peaked.

[12]This number is a result of the formula $\sum_{k=1}^{m} \binom{m-1}{k-1} = 2^{m-1}$, where we can identify the index k with the ranking position of the peak and the binomial coefficient as the quantity of possibilities to order the remaining parties. For a detailed proof see Escoffier et al. (2008).

The answer to this question depends on several factors, e.g. on the distribution of preferences and the number of individuals in the profile. With a growing number of individuals, the probability of getting a single-peaked profile decreases. Given a uniform distribution of preferences, Escoffier et al. (2008) give a formula for this probability. For $n = 250$, this probability is almost zero, which justifies the explained method.

9.2 Method: Generalized single-peaked preferences over connected coalitions and small coalitions

To check whether a profile satisfies property *con* or *small* more elaborate methods are necessary. As the properties are very similar, only *con* is explained here. *small* works analogously, see the remarks on isomorphism in Section 7.2.2.

For a given ordering of political parties, the set of generalized single-peaked preferences on the set of connected coalitions is enormous. Hence, a check that compares the list of datasets with the reference set is impossible. Therefore, we implemented an algorithm with the object-oriented programming language "Java". A pseudo-code is given in Algorithm 1. The algorithm uses our data, which is saved in an array. It generates all 60 possible orderings of the five political parties. The output variables are two counters, to be precise vectors of dimension 60. The first one is *gsp_counter*, which counts for each ordering the number of preference rankings which are generalized single-peaked with respect to this ordering. The second one is *step_counter*, which counts for each ordering the number of marked coalitions before the algorithm stops.

Example 5. We explain the algorithm based on a concrete example with $j = 2$ and Ordering[2]= $LGSUF$. First, all coalitions, which are not connected with respect to this ordering are deleted. Recall that given one concrete ordering, only 15 out of 31 coalitions are connected. With the remaining coalitions, the graph in Figure 9.1 is constructed.

Now, each individual preference ranking has to be checked one after another. In our example, we check only individual i, whose ranking starts with the sequence $(GS, S, G, SU, \cancel{GU}, GSU, \dots)$. Note that GU is crossed out as it is not connected with respect to the ordering and therefore was already deleted in the first step.

At this point, the *while*-loop of the algorithm begins:

- We start with $k = 1$ and mark the first coalition, i.e. the peak GS. Trivially, every shortest path to GS is marked.

```
Data: Preference profile (≿_1, ..., ≿_n)
Result: Number of generalized single-peaked preferences per ordering.
generate Ordering[60];
for j = 1 to 60 do
    gsp_counter[j] = 0;
    step_counter[j] = 0;
    use Ordering[j];
    delete irrelevant coalitions;
    construct Graph[j];
    for i = 1 to n do
        use reduced_preference (≿_i = C^i_1, ..., C^i_15);
        stop = false;
        k=0;
        while stop = false do
            k++;
            mark coalition C^i_k;
            if A shortest path between two marked coalitions exists which uses
            a non-marked coalition then                    /* ≿_i is not gsp */
                step_counter[j] = step_counter[j] + (k-1);
                stop=true;
            end
            if k=15 then                                   /* ≿_i is gsp */
                gsp_counter[j]++;
                step_counter[j] = step_counter[j] + k;
                stop=true;
            end
        end
        i++;
    end
    j++;
end
return gsp_counter, step_counter;
```

Algorithm 1: Generalized single-peakedness check

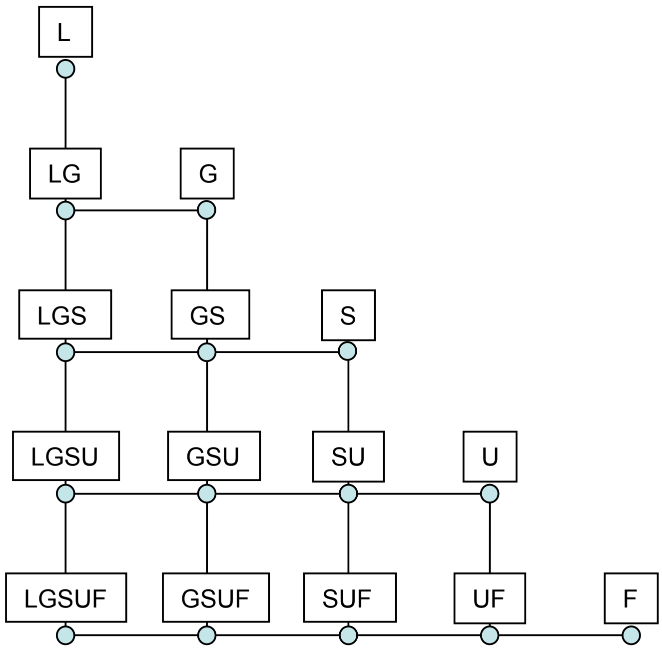


Figure 9.1: The graph of connected coalitions for the ordering $LGSUF$

- In the next iteration loop ($k = 2$), we mark coalition S. It is a direct neighbor of GS and hence lies on the shortest path.
- For $k = 3$, we mark the third coalition G, which is also a neighbor of GS. There is only one shortest path from G to S, namely (G, GS, S) which contains only marked coalitions.
- In the fourth step ($k = 4$), we mark SU. It is a neighbor of S, which is marked. However, there exists a shortest path (SU, GSU, GS) that uses the non-marked coalition GSU. Hence, the while-loop stops at this point.

The considered ranking $\succsim_i$ is *not* generalized single-peaked with respect to the ordering $LGSUF$. The algorithm stops and we increase the step counter ($step\_counter[2]$) of the particular ordering ($j = 2$) by $k - 1 = 3$, as three coalitions have been marked. The counter $gsp\_counter[2]$, which counts the number of preferences that are generalized single-peaked with respect to this ordering, is not increased.

Then, the next individual is analyzed. In our example, individual $i + 1$ has the following ranking which is already reduced to the connected coalitions.

$$(U, UF, SU, S, F, SUF, GSU, GS, G, GSUF, LGSU, LGSUF, LGS, LG, L)$$

It is easy to verify, that here the algorithm stops not until $k = 15$, as the preference is generalized single-peaked with respect to $LGSUF$. Hence, the $step\_counter[2]$ is increased by 15 and the $gsp\_counter[2]$ is increased by one.

N° parties	N° con. coalitions	N° preferences	N° GSP preferences	ratio
1	1	1	1	1
2	3	6	4	.666667
3	6	720	56	.077778
4	10	3628899	5204	.001434
5	15	$1.3 \cdot 10^{12}$	5030332	.000004

Table 9.1: Ratio of generalized single-peaked preferences

As soon as every individual preference ranking has been checked, the next ordering of political parties ($j++$) is taken to construct the next graph and so on until each of the 60 orderings has been checked.

The algorithm returns two counters, *gsp_counter* and *step_counter*, where the numbers for each ordering are saved.

The explained algorithm works exactly and counts all preferences that are generalized single-peaked for coalitions that are connected, i.e. have property *con*. However, the restriction given by *con* is very strong, i.e. only few preferences fit exactly in this structure.

Therefore, we approximate generalized single-peakedness by allowing up to four transpositions t in the ranking. A *transposition* of a preference ranking is a permutation of exactly two coalitions.[13] Note that computing time of the algorithm grows exponentially in the number of transpositions. A preference ranking is *almost generalized single-peaked* if up to four transpositions are necessary to transform it into a generalized single-peaked preference ranking. For $t = 4$, the algorithm ran approximately 110 minutes per survey. For purposes of readability, a more detailed description of the implementation is skipped here.[14]

9.3 Results expected in an impartial culture

Which results do we expect when we run the algorithm with our data? To evaluate whether the results are in a sense good or unexpected, we try to answer this questions in two different ways. First, the analytical one, where we calculate the number of single-peaked preferences in relation to the overall possibilities of preferences given one fixed order. Results are given in Table 9.1.

[13] For instance, the ranking (1-2-3-4) can be transposed (3-2-1-4); and (3-2-1-4) can be transposed into (3-1-2-4).

[14] More information about the implementation with Java is available on request.

For us, the important is the last row: With five parties and 15 connected coalitions, a randomly picked preference ranking satisfies generalized single-peakedness with a probability very close to zero (0.0004%). Therefore, the second way to answer the question is given by data simulation. We considered 20 samples of 250 random datasets which were generated by uniformly distributed data. Then, we tested each of the samples on generalized single-peakedness. The result was really surprising: For each ordering, and each sample, the number of generalized single-peaked preferences was zero. The result is the same for connected and small coalitions. So, in an impartial culture, we expect that we will not find generalized single-peaked preferences no matter which ordering we consider.

10 Results

The result section is structured as follows: First, we analyze plurality voting over parties and over coalitions. Then, we state the results on generalized single-peakedness with the properties *single*, *con* and *small*. Finally, the influence of coalition size and of extreme parties is considered.

10.1 Distribution of preferences: peak-coalitions

In this subsection we present the distribution of the peaks of the preferences from the participants of the survey. In the first paragraph we have a closer look at the singletons, in the second paragraph we consider all coalitions.

10.1.1 Most preferred singletons

For politicians, the most relevant question is: *Which party wins the election?*

Assume a hypothetical election, where each of our respondents votes for the political party he likes best, i.e. which occurs first in his individual ranking list as a single party coalition.[15] Table 10.1 shows the result of this hypothetical plurality-voting for the five parties.

Party	S_1	S_2
Union	45.5 %	52.4 %
SPD	25.9 %	18.4 %
FDP	16.2 %	14.4 %
Green	11.3 %	13.2 %
Left	1.1 %	1.6 %

Table 10.1: Hypothetical plurality voting for parties, $n_1 = 266$, $n_2 = 250$

[15]For instance, a ranking list (SPD-Green, Green, SPD,. . .) would be counted as a vote for the Green party and a list (SPD-Green, SPD, Green,. . .) or (SPD, Green, SPD-Green,. . .) would be counted for the SPD.

Although the percentage varies in the surveys, the ranking of the parties is the same for both surveys. In both surveys, “Union” is the favorite party with 45.5% and 52.4% respectively, followed by “SPD”, “FDP” and “Green”. The “Left” is least preferred with less than two percent of the votes.

Plurality voting is motivated by the German federal election system where every voter has two votes. The first vote decides which candidate is sent directly from a constituency to Parliament. The second vote is more important in the sense that it determines the proportional representation of each party in Parliament. It is possible to split votes for the two ballots but this has no effect on actual coalition formation. Parties are not bounded to agreements they have made before elections as coalition formation depends on final election results. Even though most of the governing coalitions consist of exactly two political parties, every voter has to decide for only one party for his second vote. This corresponds to the plurality rule we have just analyzed.

To give a measure of substantive agreement, the Herfindahl-Hirschman-Concentration-Index H (see Hirschman (1964)) is calculated.[16] We therefore sum up the squares of the numbers in Table 10.1. For the first survey, we obtain $H_1^P = 0.3468$ (normalized: 0.1835) and for the second survey $H_2^P = 0.3132$ (normalized: 0.1415). If all individuals preferred the same party, the index would be $H^P = 1$ (normalized: 1).[17] In case of a uniform distribution, it would be $H^P = \frac{1}{m} = \frac{1}{5} = 0.2$ (normalized: 0).

Note that the restriction of reducing preferences to the ranking of singletons is strong. A dataset where the five singletons were ranked on positions 1 to 5 would give the same information as a dataset with singletons ranked on positions 27 to 31 as long as the ordering on the subset of singletons remained the same.

10.1.2 Most preferred coalitions

Using plurality voting, we elicit the peaks of the preferences over (all) coalitions. In contrast to plurality voting over singletons, this voting mechanism has no equivalent in the German voting system. To answer the question *which coalition wins our hypothetical election?*, i.e. which coalitions are most frequently ranked first, we consider the first places of data for both surveys S_1 and S_2 as illustrated in Table 10.2.

We obtain the result that the coalition of Union and FDP tops the list in both surveys, as it achieves 23.5% of the votes in S_1 and 25.9% in S_2. Note that this coalition was the government coalition during the second survey S_2, but not when

[16] The inverse of the Herfindahl-Hirschman-Index is also known as Laakso-Taagepera index (see Laakso and Taagepera (1979), Taagepera and Grofman (1981))

[17] In H^P the ^{P} stands for 'party' in contrast to ^{C} for 'coalition' in the following paragraph.

Coalition	S_1	S_2
Union, FDP	23.3 %	26.8 %
Union	18.8 %	15.2 %
SPD, Green	14.7 %	10.4 %
FDP	7.9 %	4.8 %
SPD	7.1 %	8.0 %
Union, Green	5.6 %	8.4 %
SPD, FDP	4.5 %	2.0 %
Union, SPD	4.1 %	6.8 %
Green	4.1 %	5.2 %
Union, FDP, Green	1.9 %	3.6 %

Table 10.2: Hypothetical plurality voting for coalitions, $n_1 = 266$, $n_2 = 250$; the coalitions are listed with decreasing percentages with respect to S_1

the first survey was conducted. On the second place, there is the single-party Union followed by the coalition of SPD and Green. The coalition Union and SPD, which was the government coalition during the first survey S_1, obtains 3.6% and 6.8% of the votes, which leads to positions 8 and 6, respectively. The coalitions are listed with decreasing percentages with respect to S_1. Here, for instance, FDP has a larger support than SPD. In contrast, in S_2 the party SPD is ranked first more often than FDP. We compare the implicit ordering of single parties with the result of Table 10.1 and obtain a consistent ordering for survey S_2 (Union-SPD-FDP-Green-Left) but an inconsistent result for the first survey due to the reverse ordering of FDP and SPD. Coalitions not listed in Table 10.2 achieved less then 2% of the votes in both surveys.

Historically, the coalition Union-FDP formed the government nine times, and SPD-Green twice (1998-2005). Except for the years 1960-1961 where the Union solely formed the government, always two parties were required (and sufficient) to achieve a majority.[18]

As in the previous paragraph, we calculate the Herfindahl-Hirschman concentration index. For the first survey, we get $H_1^C = 0.1323$ (normalized: 0.1033) and for the second survey a similar value of $H_2^C = 0.1315$ (normalized: 0.1025). If all individuals preferred the same coalition, the index would be $H^C = 1$ and in case of a uniform distribution it would be $H^C = \frac{1}{2^m-1} = \frac{1}{31} = 0.0322$. Hence, we observe a concentration on some particular coalitions, i.e. the data are biased.

[18] An overview over Germans Governing Parties is available on `http://de.wikipedia.org/wiki/Datei:German_parliamentary_elections_diagram_de.png`, retrieved December 2013.

10.2 Single-peakedness of singletons

In this paragraph we ask with respect to which ordering of political parties most of the preferences are single-peaked, i.e. are generalized single-peaked with property *single*. For every ordering of political parties the proportion of single-peaked preference-rankings is calculated, i.e. for each of the 60 orderings, the corresponding 16 preference rankings[19] are determined that are single-peaked with respect to this ordering as explained in Section 9.1. Then, the number of those preferences is summed up and divided by the total number of individuals ($n_1 = 266, n_2 = 255$). This measure for single-peakedness was also introduced in List et al. (2013) as *proximity to single-peakedness*, which equals the size of a largest subset of sample members whose combination of preferences is single-peaked divided by the overall sample size. The corresponding order of political parties is called a *largest structuring dimension*. The lower bound for the proximity to single-peakedness is given by $\frac{2^{m-1}}{m!}$ which equals $\frac{2}{15}$ for $m = 5$. This would be the case when we would consider an impartial culture. Tseltin, Regenwetter and Grofman show that the impartial culture maximizes the probability of majority cycles (see Tsetlin et al. (2003)). Hence, this assumption always gives us a lower bound of the expected largest structuring dimension.

All orderings with values higher than 20% are listed in Table 10.3 in a descending order with respect to S_1.[20] In the following, we will focus on three special orderings.

$$O_1 := (\text{LSGUF}), O_2 := (\text{LGSUF}), \text{ and } O_3 := (\text{LGFUS})$$

The ordering O_1, i.e. Left - SPD - Green -Union - FDP, coincides with the seating arrangement in the German Parliament from 1998 to 2013. However, when we look at Table 10.3, we see that the ordering with the highest percentage of single-peaked preferences is Left - Green - SPD - Union - FDP. In the first survey, 70.7% and in the second survey 62.4% of the respondents have single-peaked preferences with respect to this ordering. Hence, we highlight this ordering as $O_2 :=$ (LGSUF). Interestingly, this ordering coincides with the spectrum projected on the "Communism vs. Neo-Liberalism" axis, suggested from the "Political Compass" after evaluating the elections in 2005 and 2013.[21] O_3 will be motivated in the following sections.

Note that in O_2 and O_1 only the positions of "SPD" and "Green" are switched, but only 21.1% and 27.6% of the respondents have single-peaked preferences according to O_1.

[19] Recall Appendix C for combinatorial remarks.

[20] A detailed list is available from the author on request.

[21] see `www.politicalcompass.org/germany2005` and `http://www.politicalcompass.org/germany2013`, retrieved December 2013.

Ordering	S_1	S_2
LGSUF	70.7 %	62.4 %
LFUSG	66.9 %	58.4 %
LSUFG	54.9 %	55.2 %
LGFUS	53.0 %	51.2 %
LSFUG	44.0 %	40.4 %
LGUFS	42.1 %	38.4 %
LGSFU	41.0 %	34.0 %
LUFSG	40.6 %	31.6 %
LFSUG	36.8 %	31.6 %
LGUSF	36.1 %	32.4 %
LUSGF	29.3 %	26.8 %
LFGSU	27.4 %	
LSGUF	21.1 %	27.6 %
LSUGF	21.1 %	22.0 %
LUSFG	20.7 %	
LFUGS		22.0 %
others	< 20 %	

Table 10.3: Percentage of single-peaked preferences for a given order, $n_1 = 266$, $n_2 = 250$; listed decreasingly in S_1

The Left party is extreme in the sense that orders with this party as an extreme element are more likely to obtain higher percentages of single-peaked preferences. Orderings where "The Left" is not an extremal element have at most 6.0% of single-peaked preferences. The members of the coalition Union-FDP which formed the government during the second survey are neighboring parties in each of the top eight orderings. The former government Union-SPD (2005-2009) are neighboring parties in each of the top five orderings.

In 2013, the initial coalition negotiations on a federal level of Union and Green failed. According to our table, these two parties are neighbors not until the fifth row. The next paragraph will focus on connectedness of coalitions, i.e. coalitions that satisfy property *con*.

10.3 Generalized single-peakedness over connected coalitions

In this section, we analyze whether our datasets represent generalized single-peaked preferences over connected coalitions, i.e. property *con* according to the method explained in Section 9.2. The aim is to find the ordering which maximizes the

con Ordering	survey 1 (269) t=0	t=4	$\%t = 4$	survey 2 (250) t=0	t=4	$\%t = 4$	both surveys (519) t=0	t=4	$\%t = 4$
GLFSU	0	62	23%	0	58	23%	0	120	23%
GLFUS	0	106	39%	0	106	42%	0	212	41%
GLSFU	0	76	28%	0	63	25%	0	139	27%
GLSUF	0	117	43%	0	117	47%	0	234	45%
GLUFS	0	111	41%	0	110	44%	0	221	43%
GLUSF	0	108	40%	0	102	41%	0	210	40%
GFLSU	0	18	7%	0	18	7%	0	36	7%
GFLUS	0	8	3%	0	20	8%	0	28	5%
GFSLU	0	59	22%	0	55	22%	0	114	22%
GFSUL	0	192	71%	0	172	69%	0	364	70%
GFULS	1	109	41%	0	117	47%	1	226	44%
GFUSL	3	211	78%	5	193	77%	8	404	78%
GSLFU	0	22	8%	0	27	11%	0	49	9%
GSLUF	0	22	8%	0	32	13%	0	54	10%
GSFLU	0	59	22%	0	56	22%	0	115	22%
GSFUL	1	197	73%	1	167	67%	2	364	70%
GSULF	0	116	43%	0	125	50%	0	241	46%
GSUFL	3	210	78%	4	195	78%	7	405	78%
GULFS	0	10	4%	0	22	9%	0	32	6%
GULSF	0	19	7%	0	28	11%	0	47	9%
GUFLS	1	97	36%	0	103	41%	1	200	39%
GUFSL	4	195	72%	2	174	70%	6	369	71%
GUSLF	0	92	34%	0	89	36%	0	181	35%
GUSFL	1	186	69%	0	183	73%	1	369	71%
LGFSU	0	193	72%	0	162	65%	0	355	68%
LGFUS	8	205	76%	12	171	68%	20	376	72%
LGSFU	2	206	77%	1	171	68%	3	377	73%
LGSUF	8	215	80%	4	204	82%	12	419	81%
LGUFS	11	202	75%	6	180	72%	17	382	74%
LGUSF	2	203	75%	1	195	78%	3	398	77%
LFGSU	0	155	58%	0	136	54%	0	291	56%
LFGUS	0	185	69%	2	158	63%	2	343	66%
LFSGU	1	147	55%	0	140	56%	1	287	55%
LFUGS	1	198	74%	1	186	74%	2	384	74%
LSGFU	1	164	61%	0	144	58%	1	308	59%
LSGUF	1	183	68%	0	181	72%	1	364	70%
LSFGU	1	154	57%	0	138	55%	1	292	56%
LSUGF	1	195	72%	0	196	78%	1	391	75%
LUGFS	1	158	59%	0	139	56%	1	297	57%
LUGSF	1	157	58%	0	148	59%	1	305	59%
LUFGS	1	166	62%	0	159	64%	1	325	63%
LUSGF	2	175	65%	3	158	63%	5	333	64%
FGLSU	0	26	10%	0	38	15%	0	64	12%
FGLUS	0	25	9%	0	22	9%	0	47	9%
FGSLU	2	80	30%	2	75	30%	4	155	30%
FGULS	2	66	25%	0	65	26%	2	131	25%
FLGSU	0	65	24%	0	64	26%	0	129	25%
FLGUS	0	58	22%	0	70	28%	0	128	25%
FLSGU	1	70	26%	1	67	27%	2	137	26%
FLUGS	1	64	24%	0	82	33%	1	146	28%
FSGLU	2	64	24%	0	65	26%	2	129	25%
FSLGU	1	23	9%	0	34	14%	1	57	11%
FUGLS	2	70	26%	1	66	26%	3	136	26%
FULGS	1	27	10%	1	22	9%	2	49	9%
SGLFU	0	17	6%	0	20	8%	0	37	7%
SGFLU	0	33	12%	0	42	17%	0	75	14%
SLGFU	1	63	23%	0	46	18%	1	109	21%
SLFGU	1	32	12%	0	44	18%	1	76	15%
SFGLU	0	33	12%	0	46	18%	0	79	15%
SFLGU	0	11	4%	0	18	7%	0	29	6%

Table 10.4: Generalized single-peaked preferences over connected coalitions: Share of preferences satisfying property *con* for all orderings

share of generalized single-peaked preferences among the respondents of our survey considering only connected coalitions. Recall that depending on the ordering of political parties, the set of connected coalitions contains different elements.

Table 10.4 gives an overview over both surveys with no transpositions ($t = 0$) and up to four transpositions ($t = 4$). The first column indicates the underlying ordering in the corresponding row. There exist 60 different orderings which are all checked by the algorithm. Each of the following columns is divided in three sub-columns, where the first one indicates the number of preferences which are generalized single-peaked without transpositions, the second indicates the number with up to four transpositions and the last one shows the corresponding percentage for $t = 4$. Data are grouped for Survey 1 and Survey 2 and for the aggregated datasets. In each column the orderings O_1, O_2 and O_3 are highlighted.

More details about the step-counter for each survey and each number of transpositions can be found in the Appendix (see Table E.1 and Table E.2).

Without transpositions, $O_3 = LGFUS$ maximizes the number of generalized single-peaked preferences. For this ordering, 20 of 519 (i.e. 3.8%) of all preferences are generalized single-peaked. Due to the reasons already mentioned in the theoretical part, this small percentage is not surprising. The reason why the algorithm stopped so early was that often the alternative ranked on the second position was not a direct neighbor of the peak, i.e. there exists no link between the coalitions in the corresponding graph (recall Figure 7.3). For instance, if an individual ranking starts with two singletons ($U, F, UF, \dots$) it is never generalized single-peaked independently of the ordering. This is due to the fact, that two connected coalitions of the same size (i.e. the same number of parties) are never linked by an edge in our model.

To solve this problem we consider preferences which are *almost* generalized single-peaked, i.e. we allowed up to $t = 4$ transpositions. If we allow up to four transpositions, 72% of the preferences are almost generalized single-peaked with respect to O_3 and even 81% are almost generalized single-peaked with respect to $LGSUF$. This ordering is O_2, already known from Section 10.2 as the ordering which is also the best fitting for single-peaked preferences over singletons ($M = single$). Recall Figure 9.1 for the graphical interpretation.

Given the ordering of the German Parliament O_1, only one person has generalized single-peaked preferences without transpositions. With up to four transpositions, 70% of the preferences are almost generalized single-peaked with respect to O_1. Here, we see that almost generalized single-peakedness is a much weaker restriction than generalized single-peakedness.

According to our definition, the seating arrangement of the German Parliament does not seem to represent a good approximation of the perceived left-right-spectrum suggested by our survey data.

10.4 Generalized single-peakedness over small coalitions

When analyzing the preferred coalitions, we have seen that many students rank coalitions with at most two coalition partners first. Hence, we modify our model, such that all "small" coalitions are taken into account. Again, we ask the question which underlying ordering of political parties maximizes the share of generalized single-peaked preference rankings. In contrast to the algorithm used in the previous section, the set of alternatives does not change when considering different orderings. It always contains the same 15 elements, namely the five single parties and the ten coalitions consisting of two parties.

Table 10.5 shows the results. The structure of the table is similar to the one in the previous section. Again, the orderings O_1, O_2, and O_3 are highlighted. For the aggregated data, $O_3 = LGFUS$ maximizes the share of generalized single-peaked preferences when no transpositions are allowed. Here, 28 of 519 are generalized single-peaked for small parties. When considering up to four transpositions, we obtain 425, i.e. 82%, of almost generalized single-peaked preferences with respect to the ordering $O_2 = LGSUF$.

The German Parliament seating arrangement O_1 performs poorly in this model. Only two rankings satisfy generalized single-peakedness with respect to this ordering.

We conclude that when no transpositions are allowed, O_3 maximizes the percentage of generalized single-peaked preferences over connected coalitions as well as over small coalitions. When considering up to four transpositions, O_2 performs even better for our data. This ordering O_2 also maximizes the share of single-peaked preferences over single-parties. As an additional result, the ranking of the hypothetical plurality voting (see Table 10.1) is consistent with O_2.

10.5 The size of coalitions

In this section we describe how the size of a coalition, i.e. the number of contained parties influences its valuation. Figure 10.1 shows for each of the 31 ranking positions how often a coalition of a given size is ranked in that position.

On the first rank, i.e. the favorite coalition, 287 of 519 participants place a two-party coalition, 190 place a single-party and 38 place a coalition consisting of three parties. On the last rank, i.e. the worst coalition, 342 participants rank a single party whereas 107 place the coalition containing all five parties. Coalitions containing three or four members are not very popular in the beginning and at the end of the ranking but

small	survey 1 (269)			survey 2 (250)			both surveys (519)		
Ordering	t=0	t=4	%$t=4$	t=0	t=4	%$t=4$	t=0	t=4	%$t=4$
GLFSU	0	32	12%	0	24	10%	0	56	11%
GLFUS	0	53	20%	0	58	23%	0	111	21%
GLSFU	0	27	10%	0	35	14%	0	62	12%
GLSUF	0	41	15%	0	52	21%	0	93	18%
GLUFS	0	69	26%	0	61	24%	0	130	25%
GLUSF	0	46	17%	0	61	24%	0	107	21%
GFLSU	0	29	11%	0	35	14%	0	64	12%
GFLUS	0	44	16%	0	52	21%	0	96	18%
GFSLU	0	39	14%	0	48	19%	0	87	17%
GFSUL	0	201	75%	0	175	70%	0	376	72%
GFULS	0	62	23%	0	64	26%	0	126	24%
GFUSL	16	215	80%	11	199	80%	27	414	80%
GSLFU	0	22	8%	0	33	13%	0	55	11%
GSLUF	0	42	16%	0	46	18%	0	88	17%
GSFLU	0	34	13%	0	49	20%	0	83	16%
GSFUL	3	192	71%	2	180	72%	5	372	72%
GSULF	0	41	15%	0	54	22%	0	95	18%
GSUFL	9	218	81%	6	197	79%	15	415	80%
GULFS	0	33	12%	0	46	18%	0	79	15%
GULSF	0	31	12%	0	49	20%	0	80	15%
GUFLS	0	43	16%	0	46	18%	0	89	17%
GUFSL	12	204	76%	9	188	75%	21	392	76%
GUSLF	0	31	12%	0	46	18%	0	77	15%
GUSFL	3	199	74%	1	194	78%	4	393	76%
LGFSU	0	192	71%	1	162	65%	1	354	68%
LGFUS	18	215	80%	10	182	73%	28	397	76%
LGSFU	7	192	71%	3	174	70%	10	366	71%
LGSUF	9	222	83%	6	203	81%	15	425	82%
LGUFS	15	208	77%	7	190	76%	22	398	77%
LGUSF	2	208	77%	2	188	75%	4	396	76%
LFGSU	1	148	55%	2	145	58%	3	293	56%
LFGUS	0	188	70%	1	175	70%	1	363	70%
LFSGU	1	169	63%	2	146	58%	3	315	61%
LFUGS	0	183	68%	1	165	66%	1	348	67%
LSGFU	3	152	57%	1	136	54%	4	288	55%
LSGUF	0	182	68%	2	182	73%	2	364	70%
LSFGU	0	149	55%	0	149	60%	0	298	57%
LSUGF	0	194	72%	2	185	74%	2	379	73%
LUGFS	2	158	59%	0	143	57%	2	301	58%
LUGSF	2	162	60%	2	150	60%	4	312	60%
LUFGS	0	162	60%	2	148	59%	2	310	60%
LUSGF	2	177	66%	2	153	61%	4	330	64%
FGLSU	0	51	19%	0	55	22%	0	106	20%
FGLUS	0	37	14%	0	47	19%	0	84	16%
FGSLU	2	57	21%	1	60	24%	3	117	23%
FGULS	0	40	15%	0	44	18%	0	84	16%
FLGSU	0	42	16%	0	63	25%	0	105	20%
FLGUS	0	35	13%	0	37	15%	0	72	14%
FLSGU	1	44	16%	1	65	26%	2	109	21%
FLUGS	0	27	10%	0	34	14%	0	61	12%
FSGLU	1	37	14%	0	49	20%	1	86	17%
FSLGU	0	33	12%	0	51	20%	0	84	16%
FUGLS	0	30	11%	0	35	14%	0	65	13%
FULGS	0	31	12%	0	34	14%	0	65	13%
SGLFU	0	22	8%	0	31	12%	0	53	10%
SGFLU	0	34	13%	0	35	14%	0	69	13%
SLGFU	0	17	6%	0	30	12%	0	47	9%
SLFGU	0	26	10%	0	30	12%	0	56	11%
SFGLU	0	23	9%	0	40	16%	0	63	12%
SFLGU	0	20	7%	0	41	16%	0	61	12%

Table 10.5: Generalized single-peaked preferences over small coalitions: Share of preferences satisfying property *small* for all orderings

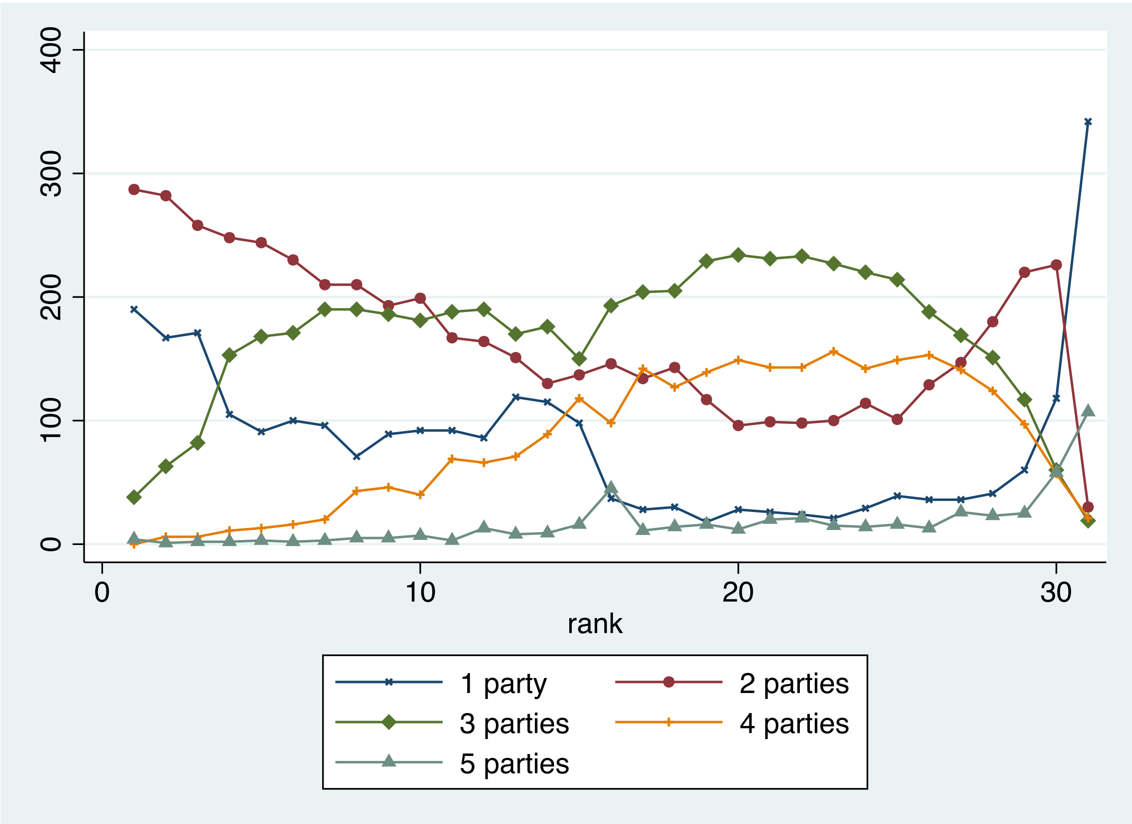


Figure 10.1: Distribution of coalitions of different sizes among ranking positions

are placed more often in the middle: most three-and four-party-coalitions are placed on ranking positions 20 and 23, respectively.

Note that in each ranking there is only one coalition with five members, but five times more coalitions with one and four parties, and even ten coalitions with two or three parties. One explanation for the number is the following. Participants divide the set of all coalitions into three groups. Their favorite party or parties are put into the first group. Coalitions which are still acceptable and preferred to those coalitions which are not able to make decisions due to their size are put into the second group. Finally, the third group contains all coalitions which include the party which they absolutely dislike to be part of the government. We will come to this effect in the following section.

10.6 The influence of one extreme political party on the preferences

In the set of political parties, the Left party takes a special position. Only 1.12% of the respondents consider it as their favorite party (see Table 10.1).[22] To get an intuition of what this means for the individual rankings, Figure 10.2 presents the survey data in a special way.

We first explain how to read the diagram. Each of the five columns has a width of 31 ranking positions and corresponds to one political party, whose name is written in the caption above the column. In each column, the ranking should be read from left to right: the leftmost element represents the coalition ranked first and the rightmost element represents the coalition ranked last, i.e. on position 31. The ranking of each respondent of the first survey[23] is represented by a tiny row. Whenever a dot is plotted, this indicates that the respective party is included in the coalition which is ranked on that position. Hence, the ranking is presented five times (once per party) but in each column different ranking positions are highlighted by a dot.

Second, we interpret the diagram. The eye-catching column is the fourth as it seems to be divided into a white and a black part (of course with some irregularities). This implies that many respondents put all coalitions which do not contain the Left party in the first half of the ranking and all coalitions containing the Left party in the second half. We do not observe this for any other political party.

[22] In the German parliamentary elections in 2005 the party received 8.7% of votes.

[23] As data for both surveys look similar, the figure for the second survey is deferred to Appendix F.1.

To verify this observation, we count the number of marked coalitions in the first half of the ranking, i.e. the positions $1-16$. For each party, we calculate the average over all respondents. The result is shown in Table 10.6.

party	av. markers pos.1-16	std.dev.
Union	8.86	2.06
SPD	8.54	1.55
FDP	8.05	1.51
Left	3.56	2.52
Greens	7.95	1.46

Table 10.6: Average numbers of coalitions containing each party in the first half of the ranking

If the markers were uniformly distributed, the average would have been $\frac{16^2}{31} = 8.26$. The number of the Left party is significantly smaller, which can be verified by a t-test (p-value=0.000). Hence, we confirmed that the Left party is ranked on worse ranking positions more often than other parties.

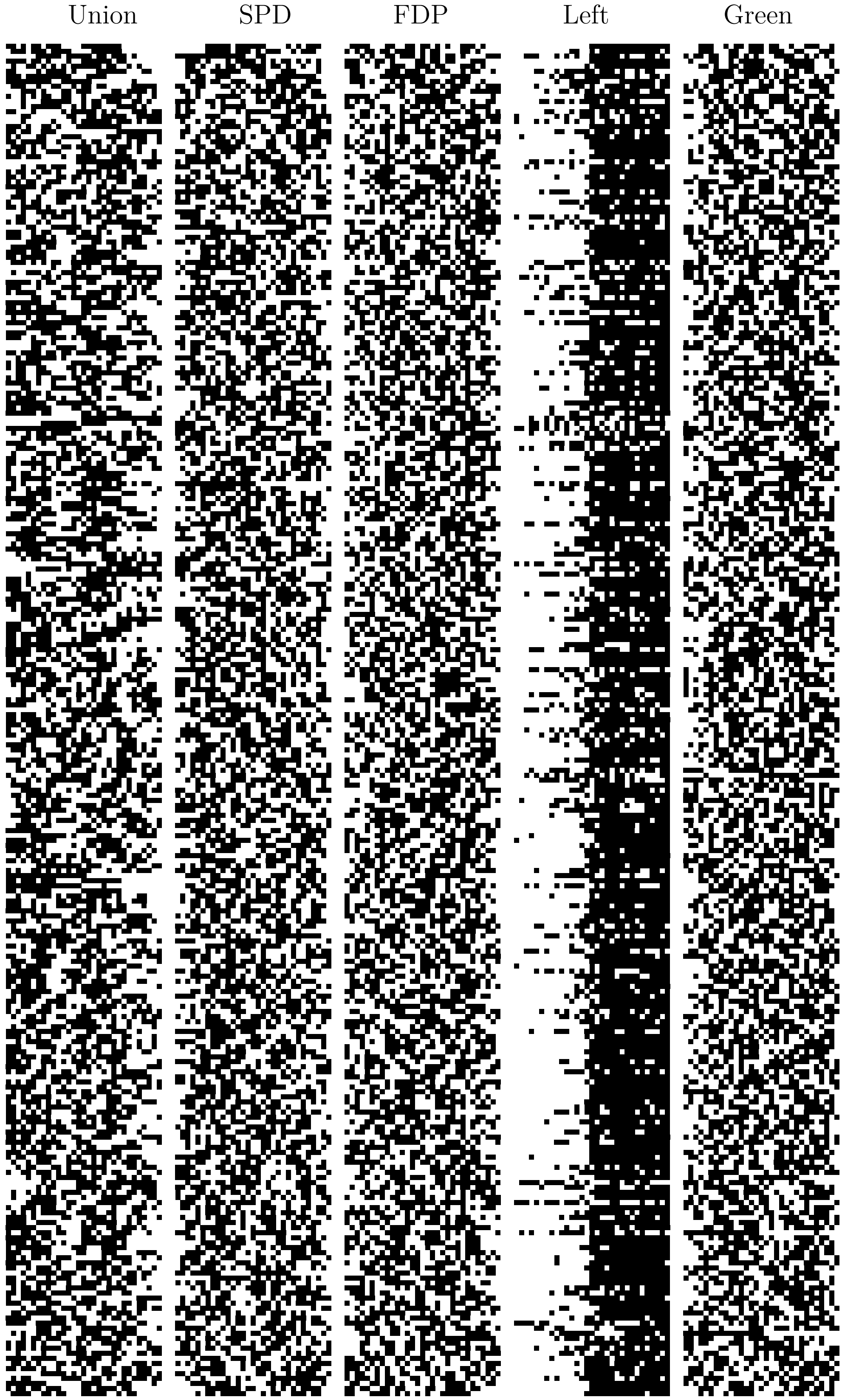


Figure 10.2: Special representation for each party: their occurrence in the individual rankings, $n_1 = 266$

11 Remarks and Conclusion

In many democracies, the concept of an election in which citizens vote for *one* political party is rarely questioned. However, many individuals do not only prefer single parties but coalitions consisting of at least two partners. Few research has been done so far to elicit the structure of these preferences. In our empirical study, we explicitly asked students to sort different coalitions in Germany according to their preferences. First, we analyzed the results of a hypothetical plurality voting over single parties and over coalitions. We verified that a majority of the respondents prefer two-party-coalitions to single parties. Second, we calculated the order of political parties O_2 which maximized the number of (one-dimensionally) single-peaked preferences over political parties. Then, we embedded the coalitions in a two-dimensional space of connected coalitions and small coalitions, respectively. This structure of generalized single-peaked preferences was already explained in the first part of this thesis. Here, we applied it to a real life example. With the help of a computer algorithm we determined the ordering of political parties O_3 which maximized the number of generalized single-peaked preferences and the ordering O_2 maximizing the number of *almost* generalized single-peaked preferences. Note that the latter ordering was the same as in the one-dimensional case.

Surprisingly, neither of these orderings coincided with the historically evolved left-right-ordering O_1 used for the seating arrangement in the German Parliament for a long time.

How shall our results be interpreted? To answer this question, we first give an overview over some limitations of our survey. With 584 respondents and 519 completely filled-in questionnaires, the result is certainly not representative for the German Society. For instance, the Left party performed very badly in our survey, whereas it achieved 11.9% of the votes in the election in 2009 and 8.6% in 2013. Hence, our results should not be compared with results in Brandenburg or Saxony-Anhalt. For instance, there may be also a greater difference between students and the working people.

Another limiting element is the huge set of alternatives. Many respondents mentioned in the fill-in question that the list of coalitions was too long to order it in a strict way. Some students explained their strategies to manage this challenge. For instance, they sorted the alternatives in two groups of which one half contained all coalitions with a particular party, e.g. the Left, and the other did not. Then, they ranked both sub-groups separately. A different strategy was to rank only the favorite coalitions without sorting the bottom half of the list.

As we asked for a ranking, we obtained only an ordinal representation of the preferences. By allowing cardinal utilities, e.g. by assigning points, two problems could have been ruled out. On the one hand, by assigning zero points, a participant could have highlighted a non-sorted list. Moreover, by assigning the same number of points, also indifferences could have been revealed.

Despite the mentioned problems, the results of the survey remain quite strong. We have shown that a large majority of the respondents prefer a coalition consisting of two parties to a coalition of one or three and more parties. One reason is that large coalitions are assumed to be incapable of making decisions due to coordination problems. A government consisting of only one party carries the risk of extreme politics. However, the voting system in Germany is still far away from voting for multiple member coalitions. It would be a challenging task to design and implement an appropriate voting system for Parliamentary elections. In our survey, we did not take into account the relative weight of a political party. In real politics, this is an important issue as, for instance, the fraction of votes determines which party provides the Chancellor.

We applied the model of generalized single-peaked preferences over connected coalitions to an example in real life and adapted it to *almost* generalized single-peaked preferences. By allowing this weakening, we obtained an ordering, such that 82% of the preferences satisfied almost generalized single-peakedness. Our algorithm can be applied to other situations and contexts as well and hence opens a novel approach to multi-dimensional voting.

In our algorithm we allow up to four transpositions. Hence, the algorithm itself is very complex and can hardly be manipulated. On the resulting graph, the multi-dimensional median rule is strategy-proof, anonymous and efficient as proven in Theorem 1.

Up to now, we assumed (generalized) single-peaked preferences either over single parties or coalitions. The third chapter of this thesis presents an experimental study of different voting rules that aggregate such preferences. Specifically, we consider the median rule, which takes the median of the announced values as social outcome and the mean rule, which takes the average of the announced values as social outcome. We implement the single-peaked preference structure on an interval by assigning specific payoff functions to the participants in the laboratory. In various sessions we test how the voting rule itself, the presence of information about others' preferences and the occurrence of manipulation costs influence individuals to vote truthfully or to manipulate strategically.

Part III

Nash Equilibrium and Manipulation in a Mean Rule Experiment

12 Motivation

Every day, people have to take decisions in small groups or committees. When it concerns money, specific thresholds or valuation in general, it can be assumed that the committee members have single-peaked preferences over the underlying alternatives. But how to find a common decision? How to aggregate votes in a "good" way? In this chapter, we focus on two popular voting rules, the mean and the median rule, and report the results of a laboratory experiment.

First, we analyze both rules in a theoretical way and describe equilibrium concepts. From the theoretical point of view, we know that the mean rule can be manipulated very easily, whereas the median rule cannot. Taking the mean value of the votes as the social outcome, this incentivizes almost every player to announce a value different from his true peak. This leads to an equilibrium with extreme individual strategies which can also be observed in the experiments. Although under the median rule, there exist infinitely many equilibria, plenty of them are inefficient. Some others can be easily ruled out when there is uncertainty over the others' preferences and behaviors and therefore uncertainty over being pivotal. Hence, the unique equilibrium which is not dominated by any of those concepts is the one where all individuals just tell the truth. Surprisingly, in our experiment we find this in half of the observations, whereas in the other half participants show a different behavior. When not being pivotal, participants tend to implement other strategies than truth-telling, i.e. we observe a variety of Nash equilibria in which participants do not announce their true peak. Therefore, we introduce manipulation costs, i.e. constant but small fees which have to be paid when deviating from the true (induced) peak. This treatment increases truth-telling as well under the median as under the mean rule.

Of course, it is difficult to find a (real-life) justification for those costs. Given that manipulation costs have to be paid, this implies that all peaks are known and this in turn implies that voting itself is not necessary as all preferences are common knowledge and therefore known by a social planner who would be able to implement an optimal social alternative. However, one simple interpretation of manipulation costs is the expected value of getting caught deviating from the peak. As an advantage of our model, stochastic effects have not to be taken into account. In general participants' decisions converge to equilibrium strategies when repeating the voting situations and inefficient equilibria do not play a crucial role in our data.

The structure of this chapter is as follows. In Section 13 we give a short overview of the literature. The model and some general notations are introduced in Section 14. We describe the median rule and state a (well-known) equilibrium in weak dominant strategies. By introducing manipulation costs, this equilibrium becomes a unique one. For the mean rule the equilibrium without costs has a nice structure but becomes more complex, or in special cases even does not exist, when costs arise. Therefore, we introduce the concept of level-k learning and summarize different properties of both rules in Section 14. We end the theoretical part by shortly analyzing welfare effects for three types of payoff functions in Section 15. Section 16 explains in detail our experimental design and in Section 17 we summarize our research questions. In Section 18, we give an overview over the observed data. Section 19 shows the statistical analysis with several parametric and non-parametric tests and a linear regression model. The focus lies on the two strategies “truth telling” and “Nash play”, but also other strategies are considered. We analyze the influence of information and framing as well as rank and manipulation costs. In Section 20 we shortly summarize the main results of this work and outline possible extensions.

13 Overview of the literature

Aggregating the votes of several decision makers has a long history in economic theory. Already Gibbard (1973) and Satterthwaite (1975) show in their famous theorem that on an unrestricted preference domain with at least three alternatives there is no strategy-proof voting rule but the dictatorship of one individual. As soon as the preference domain is restricted to single-peaked preferences on a line, possibility results are obtained as shown by Moulin (1980). Hence, from the theoretical point of view, it is very interesting to consider single-peaked preferences. This work focuses on one-dimensional preference domains. The more dimensional case, as analyzed for instance by Nehring and Puppe (2007b) or Block (2010), is not considered here.

Osborne et al. (2000) study the decision-making process when participation is costly. They analyze different *compromise* functions, i.e. voting rules, and show that only those players participate who have a large influence on the aggregated value. Translated to our context this means that only those individuals deviate, and therefore pay the manipulation costs, who are able to shift the aggregated value strong enough in the direction of their peak.

In their theoretical article, Renault and Trannoy (2005) analyze Nash equilibria under the mean rule when there are no manipulation costs. The equilibrium outcome is unique and is characterized by a median formula depending on the peaks and the voters weights. Voters tend to extreme values in the equilibrium. In Section 14.3.1 we formulate this result in a slightly different way and work out an algorithm for the discrete distribution of voters with equal weights. In particular, the rank of the voter is essentially for his equilibrium strategy. Renault and Trannoy (2005) focus on the protection of minorities which are important for groups in mean voting decisions. The influence of minorities is mainly determined by the group size and according to their theory manipulation decreases by increasing group size. Renault and Trannoy (2011) evaluate "the discrepancy between the average taste", i.e. the mean of peaks, "and the average vote", i.e. the mean of announced values. They give upper and lower bounds to show the effects of strategic behavior.

Ehlers et al. (2004) analyze the influence of voters when behaving strategically. They focus on Lipschitz continuous utility functions and show that "if there are at least five agents, the mean rule [...] is the unique anonymous and unanimous voting rule that meets a lower bound with respect to the number of agents needed to obtain threshold strategy-proofness".

Research on strategic behavior in the laboratory has been done for instance in the experimental study of Cherry and Kroll (2003) where primary elections are analyzed. Although strategic voting occurs not in large numbers, "low levels of strategic behavior can influence the election outcome".

Van der Straeten et al. (2010) conduct a voting experiment with four different election rules and "conclude that voters behave strategically as fas as strategic computations are not too demanding, in which case they rely on simple heuristics [...] or they just vote sincerely". Kube and Puppe (2009) ran another laboratory experiment. Here, Borda elections are analyzed when varying the information available. The more participants are informed, the more they try to manipulate. In contrast to other experiments, participants were asymmetrically informed. A result from Eckel and Holt (1989) that gives hope for real elections: In the laboratory, several rounds of "experience was necessary for strategic voting to occur".

The closest to our study are the experiments done by Marchese and Montefiori (2011) and the related working paper Marchese and Montefiori (2005) as they focus on the mean rule. Marchese and Montefiori (2011) ran a public good experiment, where the social choice rule selects the mean of the quantities the students voted for. A linear payoff function was used. A large share of votes was biased in the direction of the Nash equilibrium. They claim that "Strategic bias seems to be a mode of behavior that more persistently characterizes some players, while sincerity was a more intermittent way of playing". Surprisingly, "no group was able to reach the Nash equilibrium" which induced us to test this observation again by doing a similar analysis (see Section 18.1.3). As their work is strongly related to the article of Renault and Trannoy (2011), Marchese and Montefiori (2011) focus on group size. However, they found out that there is no significant influence of the group size on manipulation. In the following study, we fix group size during the entire experiment to five participants. Furthermore, by considering an odd number of individuals, the position of the median player is always uniquely detemined and can be analyzed.

Marchese and Montefiori (2005) compare manipulation under mean and median rule. Their "full information treatment" is similar to ours, but in their "no information" treatment, participants not even know the number of voters they are competing with. As payoff functions they used a quadratic function.

In our work we introduce several new aspects. First, we implement a strongly spiked utility function which is still Lipschitz-continuous. Second, we analyze how behavior changes when manipulation costs are introduced. Third, as we repeat the voting for several periods before allocating new peaks, we are able to observe learning behavior. Fourth, we implement different framings. As many people have different perceptions of truth-telling and lying (see for instance Ariely (2013)), we were curious whether the wording matters in a laboratory experiment.

14 Two rules

In this section, we briefly introduce the two rules *Mean* and *Median*. Considering the Nash equilibrium under full information, they have distinct properties. Under the mean rule, there is a tendency towards extreme values, whereas under the median rule truth-telling is a dominant strategy. For both rules, we analyze equilibria with and without manipulation costs. Under the median rule, manipulation costs imply that truth-telling becomes a strictly dominant strategy. Under the mean rule, manipulation costs may change the structure of the equilibrium or even cause the non-existence of equilibria. Furthermore, it is necessary to know the payoff function.

14.1 General notations

We begin with some general notations which are relevant for both rules.
Let $A = [0, M]$ be an interval of *feasible alternatives*. In our experiments, we set M equal to 100. The set of individuals is denoted by $I = \{1, \dots, n\}$, and every individual has a peak $x_i^* \in A$ which maximizes his utility function $u_i : A \to \mathbb{R}$. In particular, u_i is a *single-peaked* function with peak x_i^*, i.e.

$$u_i(x_i^*) \geq u_i(x) \, \forall x \in A$$

and for all $\bar{x}, \bar{y} \in A$ with $x_i^* \geq \bar{x} > \bar{y}$ or $\bar{y} > \bar{x} \geq x_i^*$, we have

$$u_i(x_i^*) \geq u_i(\bar{x}) > u_i(\bar{y}).$$

Three examples of single-peaked utility functions will be given in the following section. All considered voting rules are anonymous, i.e. it has no influence on the output who announced a particular vote. Hence, we use the ordered vector of peaks $(x_1^*, \dots, x_n^*) =: x^* \in A^n$, where $x_i^* \leq x_{i+1}^*$. We say that x^* is *strictly ordered* if the the peaks are strictly increasing, i.e. $x_i^* < x_{i+1}^*$ for $i \in \{1, \dots, n-1\}$.[24]
The mean of a vector x is calculated by $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$, where $x_i \in A$ is the choice or *vote* of individual i. For the equilibrium concept, we also need the choice x_{-i} of all

[24]In a strictly ordered vector, the index i indicates the *rank* of the individual.

other individuals $j \in I$ except i. It is defined by $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in A^{n-1}$.

For our purposes, a voting rule is a function that maps a vector of votes to a unique alternative which is the *outcome*:

$$\begin{aligned} f : A^n &\longrightarrow A \\ (x_1, \ldots, x_n) &\longmapsto f(x) \end{aligned}$$

The outcome and the individual peak x_i^* influence the payoff, which is defined by an individual payoff function.

14.1.1 Three types of payoff functions

First, we consider the payoff functions when there are no manipulation costs. The payoff function $u_i(m)$ of individual i depends only on the distance from the peak to the realized group outcome $m := f(x)$.

In the following, we will use three different payoff functions.[25]

Linear payoff functions:

$$u_i(m) = a - b|x_i^* - m| \text{ with } a \in \mathbb{R} \text{ and } b > 0 \tag{14.1}$$

Quadratic payoff functions:

$$u_i(m) = a - b(x_i^* - m)^2 \text{ with } a \in \mathbb{R} \text{ and } b > 0 \tag{14.2}$$

The special payoff function:

$$u_i(m) = 10 + \min\left(\frac{380}{|m - (x_i^* - 2)|}, \frac{380}{|m - (x_i^* + 2)|}\right) \tag{14.3}$$

The linear one (14.1) is appealing because of its simplicity. It has been used in various articles, for instance by Marchese and Montefiori (2011). The quadratic payoff function (14.2) does not need an absolute value, but is still single-peaked. Therefore, it is easy for calculations and applied for instance by Marchese and Montefiori (2005). In the remainder, a payoff function of the form (14.3) is called *special payoff*

[25] As the derivatives of the payoff functions are bounded, all of them are in particular Lipschitz-continuous. This criterium was essential for the analysis by Ehlers et al. (2004).

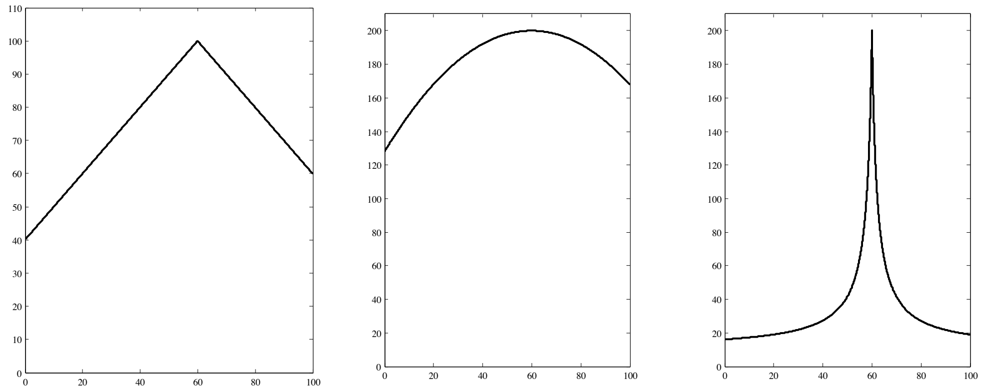


Figure 14.1: Examples for linear, quadratic and special payoff functions

function. We used this type for the experiments because of its sharp spike form. To get an intuition of the shape, examples of the three different payoff functions are visualized in Figure 14.1.

14.1.2 Payoff functions with manipulation costs

If there are manipulation costs, the structure of the payoff function looks different. The payoff function $u_i(m, x_i)$ of individual i still depends on the realized group outcome $m := f(x)$, but furthermore on the announced value x_i and manipulation costs $c(x_i, x_i^*)$. To be precise, the payoff function with manipulation costs can be written in the form

$$u_i(m, x_i) = u_i(m) - c(x_i, x_i^*)$$

were $u_i(m)$ is the payoff function without costs as defined in the previous paragraph. In the following, we will always write the extended form $u_i(m) - c(x_i, x_i^*)$ to make clear that manipulation costs occur. At this point, the only restriction we make to the cost function is $c(x_i^*, x_i^*) = 0$, i.e. the costs are zero whenever an individual announces his true peak x_i^*. In our experiment, we use *quasi-fixed* manipulation costs, i.e. $c(x_i, x_i^*) = c$ for $x_i \neq x_i^*$ and a fixed value $c > 0$. Hence, the fixed amount c has to be paid if and only if an individual deviates from his true peak x_i^*.

14.1.3 A refinement of the Nash equilibrium concept

Given a voting rule, the definition of a Nash equilibrium (Nash, 1951) reads as follows.

Definition 1 (Nash equilibrium (NE)).
$x^N = (x_1^N, \dots, x_n^N)$ is a *Nash equilibrium (NE)* if for all $i \in I$ the choice is optimal given the choice of all other individuals:

$$u_i(f(x^N)) \geq u_i(f(x_{-i}^N, x_i)) \ \forall x_i$$

A Nash equilibrium will be called *straight* Nash equilibrium (SNE), if it is a Nash equilibrium, in which all non-pivotal individuals tell the truth. Dutta and Laslier (2010) interpret this as a lexicographical preference for honesty. An individual is pivotal, if he can influence the outcome. By this manipulation, his own utility may increase or decrease.

Definition 2 (Straight Nash equilibrium (SNE)).
$x^N = (x_1^N, \dots, x_n^N)$ is a SNE, if:

a) x^N is a NE: $u_i(f(x^N)) \geq u_i(f(x_{-i}^N, x_i)) \ \forall x_i$

b) If $f(x_{-i}^N, x_i) = f(x_{-i}^N, x_i^*) \ \forall x_i$ then $x_i^N = x_i^*$.

In a Nash Equilibrium where every individual, who cannot improve his utility by manipulating, tells the truth, is called a *strong straight* Nash equilibrium.

Definition 3 (Strong straight Nash equilibrium (SSNE)).
$x^N = (x_1^N, \dots, x_n^N)$ is a SSNE, if:

a) x^N is a NE: $u_i(f(x^N)) \geq u_i(f(x_{-i}^N, x_i)) \ \forall x_i$

b) If $u_i(f(x_{-i}^N, x_i)) \leq u_i(f(x_{-i}^N, x_i^*)) \ \forall x_i$ then $x_i^N = x_i^*$.

Obviously, every strong straight Nash equilibrium is also a straight Nash equilibrium. Finally, we define a *weak straight* Nash equilibrium as a straight Nash equilibrium which is not strong.

Definition 4 (Weak straight Nash equilibrium (WSNE)).
$x^N = (x_1^N, \dots, x_n^N)$ is a WSNE, if it is a straight Nash equilibrium but not a strong straight Nash equilibrium.

For a better overview, all mentioned equilibria are visualized in Figure 14.2. Detailed examples in the context of the median rule are given in the following section.

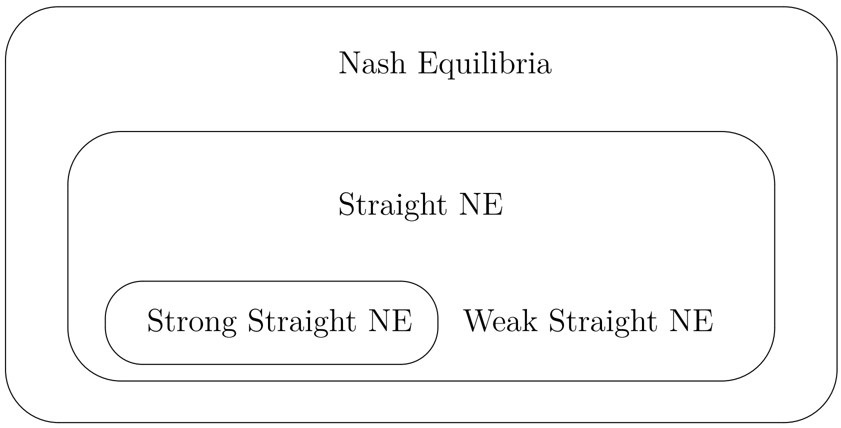


Figure 14.2: Overview over different types of Nash equilibria

14.2 Description of the median rule

One particular voting rule is the median rule which we study more closely now. From an ordered set of values, the median rule selects the element in the middle. In case of an odd number of values, this element is unique. Otherwise, the median is the average of the two middle values. We will skip the case of an even number of values also in the theoretical part, as we use a fixed number of five in our experiments. In this special case, the median is the third largest value. Hence, from here on let n be an odd number. Formally, given the values $x_1, \dots, x_n$, the corresponding ordered values are $x_{[1]}, \dots, x_{[n]}$. Thus, the median is given by

$$f^{med}(x) = x_{[\frac{n+1}{2}]}$$

14.2.1 Theoretical analysis of Nash equilibria under the median rule

Under the median rule, truth telling is an optimal strategy given that preferences are single-peaked. This result is well known as stated by Moulin (1980), but also explained quickly:

Let m be the median of the ordered vector of peaks $(x_1^*, \dots, x_n^*)$, i.e. $m = x^*_{\frac{n+1}{2}}$. For individual $\frac{n+1}{2}$ it is optimal to announce his peak $x^*_{\frac{n+1}{2}}$, as he is the median voter and therefore pivotal. If individual i with $i < \frac{n+1}{2}$ proposes a value x_i which is smaller or equal than m, the social outcome does not change at all. If he suggest a value x_i larger than m, the new median m' will be larger than m, to be precise

$m' = \min(x_i, x_{\frac{n+3}{2}})$. As $m - x_i^* < m' - x_i^*$, there is no improvement for individual i. The argumentation for an individual with a peak larger than the median is analogous. Therefore, it is an optimal strategy for each individual to propose his true peak under the median rule. To be precise, truth-telling is weakly dominant for every individual.

Theorem 3. *Under the median rule, truth-telling is the unique strong straight Nash equilibrium, i.e.* $x^N = (x_1^*, \ldots, x_n^*)$.

Proof. The proof is divided into an existence and an uniqueness part.

- **Existence.**
  Truth-telling is a SSNE. As truth-telling is a weakly dominant strategy for every individual, it is obviously a Nash equilibrium. As $x_i^N = x_i^*$ for every i, it is also a SSNE.
- **Uniqueness.**
  Each SSNE is truth-telling. Let y^N be a SSNE with $y_k^N \neq y_k^*$ for one individual k. As truth-telling is a weakly dominant strategy, it holds:

  $$u_k(f^{med}(y_{-k}^N, y_k^*)) \geq u_k(f^{med}(y_{-k}^N, y_k^N))$$

  From the definition of a SSNE, it follows that $y_k^N = y_k^*$.

□

Note that this Nash equilibrium is the only strong straight Nash equilibrium, but not unique without manipulation costs. At this point, we are able to give some examples of other Nash equilibria.

Example 6. Consider the peak distribution $x^* = (5, 10, 22, 30, 57)$. Then, there are infinitely many Nash equilibria if $A = [0, 100]$ is a continuous interval and still a large number of Nash equilibria in pure strategies if we restrict the votes to integers.

- According to Theorem 3, the only strong straight Nash equilibria is truth-telling, i.e. x^*. The outcome in the SSNE is $f^{med}(x^*) = x_3^* = 22$.
- Equilibria, in which the outcome is the same as the outcome under truth-telling, i.e. $f^{med}(x^N) = f^{med}(x^*)$, are called weak straight Nash equilibria according to Definition 4. The median voter reports his preferences truthfully. Voters with a peak smaller than the median peak announce a value smaller than $x^*_{\frac{n+1}{2}}$ and individuals with a peak which is higher than the median peak announce a value larger than $x^*_{\frac{n+1}{2}}$. In the given example, one WSNE is $x^N = (7, 10, 22, 31, 56)$. It is important, that individual 3 announces his peak 22 truthfully, but of course other individuals may announce their true

peaks as well. The equilibrium x^N is a straight Nash equilibrium, because all non-pivotal individuals announce the truth, as the set of non-pivotal individuals is empty. As there exist individuals who do not tell the truth but who cannot improve by deviating from the peak, the equilibrium is not a SSNE.

- Equilibria, in which the outcome is different to the outcome under truth-telling, i.e. $f^{med}(x^N) \neq f^{med}(x^*)$ can be either efficient or inefficient. For instance, if $\frac{n+3}{2}$ or more individuals state the same value v, this value will be selected by the median rule and cannot be changed by deviation of one individual.
  - An equilibrium which is not straight but efficient is $(10, 10, 10, 10, 10)$. With any other outcome, individual 2 will be worse off.
  - A non-straight and also inefficient equilibrium is $(0, 0, 0, 0, 0)$ as the outcome is smaller than the smallest peak $(0 < x_1^* = 5)$. In particular a value $v < x_1^*$ (or $v > x_n^*$) leads to a non-efficient outcome, as x_1^* (or x_n^*) would be a Pareto improvement. These "bad" equilibria are analyzed for instance by Cason et al. (2006), Saijo et al. (2007) and Yamamura and Kawasaki (2013).

  Obviously, such equilibria are not straight.

In the following, we eliminate all equilibria which are not SSNE by introducing manipulation costs.

14.2.2 Median rule with manipulation costs

As mentioned in the previous paragraph, truth-telling is only a weakly dominant strategy in a scenario without costs. Now, we consider the case of manipulation costs $c(x_i, x_i^*)$ which have to be payed if and only if the suggested value x_i is not equal to the peak x_i^*. With this small change, truth-telling becomes a strictly dominant strategy.

The structure of $c(x_i, x_i^*)$ is irrelevant for this argument for the median rule, as long as

$$c(x_i^*, x_i^*) < c(x_i, x_i^*) \text{ for all } x_i \neq x_i^*.$$

The simplest example which is considered in the experiments are quasi-fixed costs, i.e. $c(x_i^*, x_i^*) = 0$ and $c(x_i, x_i^*) = c$ for $x_i \neq x_i^*$ and a constant value $c \in \mathbb{R}^+$.

By introducing manipulation costs, inefficient equilibria are ruled out and the Nash equilibrium becomes unique. This fact is very important for the later analysis of the experimental results. In the following, we analyze the mean rule, for which truth-telling is hardly ever an optimal strategy.

14.3 Description of the mean rule

The mean rule, also known as average rule, computes the arithmetic mean f^{mean} of the given values $x_1, \dots, x_n$. It is defined by

$$f^{mean}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Here, f^{mean} is the social outcome and the values $x_1, \dots, x_n$ are the announced values of the participants. The mean rule is highly manipulable, i.e. only rarely it is optimal to announce the true peak. In most constellations, announcing an extreme value leads to a social outcome which is closer to the individual peak. Given that preferences are single-peaked, the Nash equilibrium can be determined by an algorithm explained in the following section.

14.3.1 Theoretical analysis of Nash equilibria under the mean rule with full information

First, we state an adapted version of the Proposition 1 of Renault and Trannoy (2005), which determines the unique equilibrium allocation. In the original version of their paper, individuals may have different weights. Here, every individual has the weight $\frac{1}{n}$.

Theorem 4 (Renault and Trannoy (2005))**.** *The average voting game has a Nash equilibrium. Furthermore, the equilibrium allocation, $\bar{x}_N$, is unique and is given by*

$$\bar{x}_N = \min\{b_{i^*}, M\frac{i^*}{n}\},$$

where $i^* = \min\{i \in I : M\frac{i}{n} \geq b_{i+1}\}$ and b_i are the peaks in a decreasing order.

If preferences are single-peaked and peaks are ordered increasingly and distinct, then the Nash equilibrium under full information is unique. This result is stated formally in Theorem 5.

Theorem 5. *Let $(u_1, \dots, u_n)$ be single-peaked functions w.r.t. $(x_1^*, \dots, x_n^*)$ and let x^* be strictly ordered. Then, there exists a unique Nash equilibrium x^N.*
The equilibrium x^N is of the form $(\underbrace{0, \dots, 0,}_{(i_0-1)-times}\ x_{i_0}^N, \underbrace{M, \dots, M}_{(n-i_0)-times})$ for some $i_0 \in I$ and $x_{i_0}^N \in [0, M]$, where M is the upper endpoint of the feasible interval.

Proof. First, we show by contradiction that a Nash equilibrium $(x_1^N, \dots, x_n^N)$ is ordered by value, i.e. $x_i^N \leq x_{i+1}^N$ for all $i \in \{1, \dots, n-1\}$.
Assume there exists a Nash equilibrium $y = (y_1, \dots, y_n)$ with $y_j > y_{j+1}$ for some j.[26]

Case 1: $\bar{y} > x_j^*$, i.e. the mean of the NE is larger than j's peak.
As $y_j > y_{j+1} \geq 0$, there exists $\varepsilon > 0$, such that $y_j - \varepsilon \in [0, M]$, (choose ε small enough, i.e. $\varepsilon < n(\bar{y} - x_j^*)$, see below). So, by choosing $y_j - \varepsilon$ instead of y_j and given all other individuals' choices y_{-j} remain the same, the utility of j changes.

$$u_j\left(\frac{1}{n}\left(y_j - \varepsilon + \sum_{i \neq j} y_i\right)\right) = u_j\left(\bar{y} - \frac{\varepsilon}{n}\right)$$

As $x_j^* < \bar{y} - \frac{\varepsilon}{n} < \bar{y}$, it follows by single-peakedness of u_j that

$$u_j\left(\bar{y} - \frac{\varepsilon}{n}\right) > u_j(\bar{y})$$

We see that j's utility increases which means that she can manipulate.

Case 2: $\bar{y} \leq x_j^*$, i.e. the mean of the NE is smaller than j's peak.
The proof is analogous to case 1. Here, individual $j+1$ can increase his utility by choosing a value $y_{j+1} + \varepsilon$ which is greater than y_{j+1}. $y_{j+1} + \varepsilon \in [0, M]$ exists, as $M \geq y_j > y_{j+1}$.

Hence, y cannot be a Nash equilibrium and we conclude that every Nash equilibrium must be of the form $(x_1, \dots, x_n)$ with $x_i < x_{i+1}$.

Second, we show that in a Nash equilibrium there can be at most one value x_{i_0} distinct from zero or M, i.e. $x_i^N \in \{0, M\}$ for all $i \backslash \{i_0\}$.
By contradiction, let us assume that in a Nash equilibrium y there exist k, l with $y_k, y_l \notin \{0, M\}$. Without loss of generality, $k < l$. Hence, $x_k^* < x_l^*$. Obviously, the mean value cannot equal both peaks, i.e. $x_k^* \neq \bar{y}$ or $x_l^* \neq \bar{y}$.
Wlog $x_k^* \neq \bar{y}$. We show that $y_k \notin \{0, M\}$ cannot be optimal.

Case 1: $\bar{y} < x_k^* < x_l^*$. As $y_k, y_l \notin \{0, M\}$, there exists $\varepsilon_k > 0$ such that:

$$u_k\left(\bar{y} + \frac{\varepsilon_k}{n}\right) > u_k(\bar{y}),$$

i.e. k can increase his utility by announcing a value $x_k = y_k + \varepsilon_k$.

Case 2: $x_k^* < x_l^* < \bar{y}$. Analogous to case 1, when l announces an appropriate value $x_l = y_l - \varepsilon_l$.

[26] This condition seems to be weak, at first. But if there exist k, l with $k < l$ and $y_k > y_l$, then there also exist some consecutive individuals $j, j+1$ with $y_j > y_{j+1}$.

Case 3: $x_k^* < \bar{y} < x_l^*$. Both individuals have an incentive to manipulate which leads to k's choice $y_k - \varepsilon_k$ and l's choice $y_l + \varepsilon_l$ until one of both reaches the endpoints of the interval A: either 0 or M.

Hence, in no case it is optimal if two or more individuals announce values distinct to the interval endpoints. $\square$

Now, we know that the Nash equilibrium must be of the form

$$x^N = (\underbrace{0, \dots, 0,}_{(i_0-1)-times}\ x_{i_0}^N, \underbrace{M, \dots, M}_{(n-i_0)-times})$$

It remains to figure out which is the individual i_0. In Theorem 6, we derive a simple algorithm associating to each individual i a specific interval $A_i \subset A$.

These intervals $A_1, \dots, A_n$ are defined as follows:

$$A_i := \left[\frac{M}{n}(n-i), \frac{M}{n}(n-i+1)\right]$$

Obviously,

$$A = \bigcup_{i=1}^{n} A_i$$

and

$$\bigcap_{i=1}^{n} A_i = \bigcup_{i=1}^{n} \left\{\frac{M}{n}(n-i+1)\right\} \neq \emptyset.$$

With these intervals, we can now state the following theorem which determines the optimal strategy depending on the rank of each individual.

Theorem 6 (Determination of the Nash Equilibrium)**.** *If every individual i chooses his x_i^N according to the following algorithm, x^N is a Nash Equilibrium.*

- *Case 1 ("$x_i^* < A_i$"): $x_i^* < \frac{M}{n}(n-i)$. Then $x_i^N = 0$.*
- *Case 2 ("$x_i^* > A_i$"): $x_i^* > \frac{M}{n}(n-i+1)$. Then $x_i^N = M$.*
- *Case 3 ("$x_i^* \in A_i$"): $\frac{M}{n}(n-i) \leq x_i^* \leq \frac{M}{n}(n-i+1)$.
Then $x_i^N = nx_i^* + M(i-n)$.*

Proof. According to Theorem 5, there exists at most one individual with a choice unequal the endpoints of the interval. So let us assume, we found i_0 and calculate his optimal choice. We know by Theorem 5 that all individuals with a smaller index choose 0, and all individuals with a higher index choose M. In the following equation i_0's peak equals the mean given these choices.

$$x^*_{i_0} = \frac{1}{n}[(i_0 - 1) \cdot 0 + x^N_{i_0} + (n - i_0) \cdot M] \tag{14.4}$$

This leads us to the optimal choice of i_0:

$$x^N_{i_0} = nx^*_{i_0} + M(i_0 - n) \tag{14.5}$$

This term is smaller than 0 if and only if $x^*_{i_0} < \frac{M}{n}(n - i_0)$ (case 1) and larger than the extreme value M if and only if $x^*_{i_0} > \frac{M}{n}(n - i_0 + 1)$ (case 2). □

Example 7. For the better understanding of Theorem 6, we give an example for $n = 5$ and the peaks $(5, 10, 22, 30, 57)$. The peaks and their position relative to the individual's interval are shown in Figure 14.3.

For individuals $1, 2$ and 3, the peak is smaller than the corresponding A_i, i.e. according to Case 1, they vote 0. Individual 5 has a peak which is larger than his interval, i.e. he votes $M = 100$ (Case 2). Case 3 holds for individual 4, whose optimal response is 50. The equilibrium allocation is shown in Figure 14.4, where the mean value 30 results.

Remark: If x^* is not strictly ordered, i.e. there exist k, l with $x^*_k = x^*_l$, the Nash equilibrium may not be unique. However, if none of the peaks lies in one of the corresponding intervals, i.e. $x^*_k \notin A_k, A_l$, the Nash equilibrium strategy is still unique; namely to announce an extreme value.

But there exist multiple NE, if for one of those individuals it holds that his peak is within his interval. Let us call this individual k with $x^*_k \in A_k$. Then, the mean of the Nash equilibrium equals both peaks, i.e.

$$\bar{x}^N = x^*_k = x^*_l$$

Hence, a vector y is a NE, if the sum of the choices of k and l remains the same, i.e. $y_k + y_l = x^N_k + x^N_l$ and all other individuals' choices remain the same, i.e. $y_i = x^N_i$ for all $i \in I \backslash \{k, l\}$. Properties of this general case can also be found in the work of Renault and Trannoy (2005).

In the previous paragraphs, it was assumed that individuals have full information about the others' peaks or at least know their own position in the ranking of individual peaks, i.e. not only x^*_i but also their own i. According to Theorem 6, it is enough to know the ranking position to determine the optimal strategy. If no information is available, the theorems are not applicable. We therefore consider the concept of level-k learning.

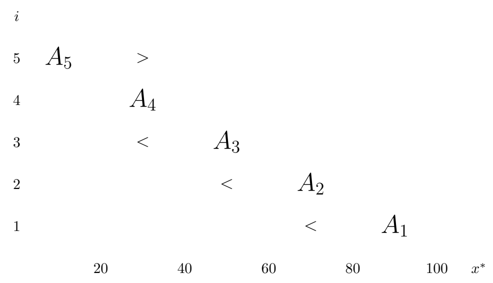


Figure 1: Ext2S4

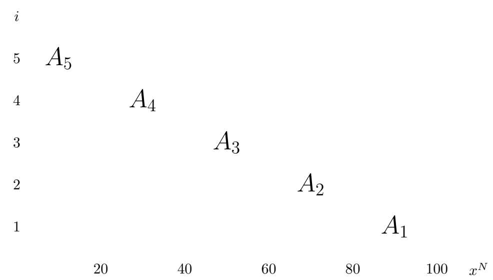
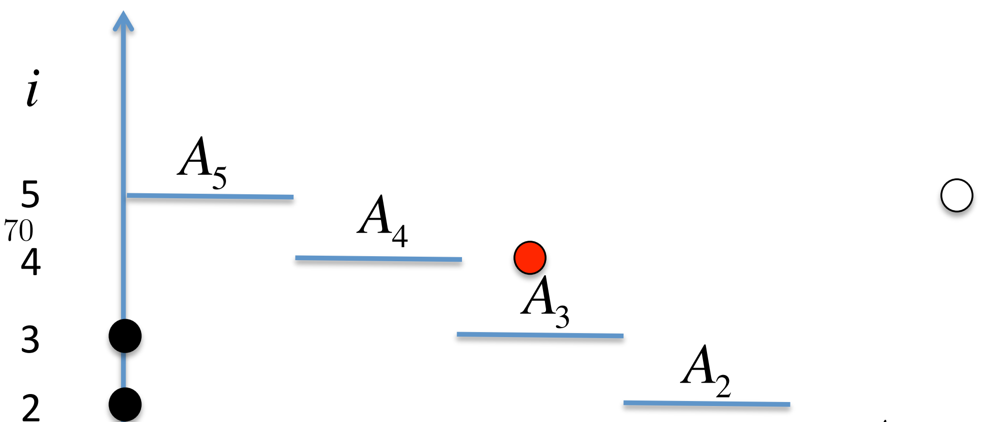


Figure 2: ExT2S9

1

Figure 14.3: Peaks $(5, 10, 22, 30, 57)$ and intervals A_i



14.3.2 Learning in the mean rule model

To analyze the behavior of individuals over periods, we also have to take into account learning effects. We consider the model of level-k learning (see Crawford and Iriberri (2007) and Stahl and Wilson (1994)).

- Level 0 player tell the truth.
- Level 1 player act optimal if all other players tell the truth.
- Level 2 player act optimal if all other players are of type level 1.
- Level k player act optimal if all other players are of type level k-1.

For mean decisions without manipulation costs, the level-k actions in our example are listed in Table 14.1. The Nash equilibrium is achieved for level-3 players at the latest.

Remark: Interestingly, in the fourth peak distribution, player 5 overshoots his true peak in the first level, but undershoots it in the following.

14.3.3 Mean rule with manipulation costs

In the model explained in the previous paragraphs, it was costless to deviate from the peak x_i^* by announcing a different value x_i. Introducing manipulation costs $c(x_i, x_i^*)$ changes the equilibrium depending on the structure of costs, the payoff function and the distribution of the peaks. We assume that each individual knows the mean value which occurs if he tells the truth. This value is denoted by m, i.e.

$$m := f^{mean}(x_{-i}, x_i^*)$$

Hence, m lies in the interval $[\frac{x_i^*}{n}, \frac{(n-1)\cdot M + x_i^*}{n}]$. Note that the assumption of knowing m is very restrictive and even in some cases with full information about peak distribution not justified. Given these assumptions, the general best response function $BR_{x_i^*}$ reads as follows.

$$BR_{x_i^*}(m) = \begin{cases} 0 & \text{if } u_i\left(m - \frac{x_i^*}{n}\right) - c(0, x_i^*) > u_i(m) \text{ and } q < 0 \\ q & \text{if } u_i(x_i^*, q) - c(q) \geq u_i(m) \text{ and } q \in [0, M] \\ M & \text{if } u_i\left(m - \frac{x_i^* + M}{n}\right) - c(M, x_i^*) > u_i(m) \text{ and } q > M \\ x_i^* & \text{otherwise} \end{cases} \tag{14.6}$$

where $q := x_i^*(n+1) - nm$. If $c = 0$, $u_i(m)$ is always smaller or equal than the utility which can be obtained by manipulation. Hence, the best response function

Peak Distribution	Player	Peak	k=0	1	2	3
1	1	50	50	90	50	50
	2	60	60	100	100	100
	3	20	20	0	0	0
	4	70	70	100	100	100
	5	10	10	0	0	0
2	1	70	70	100	100	100
	2	60	60	85	75	100
	3	50	50	25	0	0
	4	30	30	0	0	0
	5	65	65	100	100	100
3	1	10	10	0	0	0
	2	20	20	0	0	0
	3	30	30	30	0	0
	4	40	40	90	70	100
	5	50	50	100	100	100
4	1	60	60	100	100	100
	2	70	70	100	100	100
	3	10	10	0	0	0
	4	20	20	0	0	0
	5	45	45	65	25	25
5	1	70	70	100	100	100
	2	10	10	0	0	0
	3	60	60	100	100	100
	4	50	50	90	50	50
	5	20	20	0	0	0
6	1	30	30	0	0	0
	2	50	50	25	0	0
	3	70	70	100	100	100
	4	65	65	100	100	100
	5	60	60	85	75	100
7	1	40	40	90	70	100
	2	50	50	100	100	100
	3	10	10	0	0	0
	4	20	20	0	0	0
	5	30	30	30	0	0
8	1	20	20	0	0	0
	2	10	10	0	0	0
	3	45	45	65	25	25
	4	60	60	100	100	100
	5	70	70	100	100	100

Table 14.1: Level-k learning up to k=3

can be reduced to the case of "manipulation without costs":

$$BR_{x_i^*}(m) = \begin{cases} 0 & \text{if } m > \frac{n+1}{n}x_i^* \\ q & \text{if } \frac{n+1}{n}x_i^* - \frac{M}{n} \leq m \leq \frac{n+1}{n}x_i^* \\ M & \text{if } m < \frac{n+1}{n}x_i^* - \frac{M}{n} \end{cases} \tag{14.7}$$

The best response in Equation (14.6) is similar to the Nash equilibrium stated in Theorem 6 but with the additional constraint to tell the truth if deviating is too costly. As $BR_{x_i^*}$ is very sensitive to changes in the payoff function and the cost structure, a general statement is not possible. Hence, we focus on quasi-fixed costs (see Section 14.1.2) combined with the three payoff functions introduced in Section 14.1.1. Furthermore, we set $M = 100$ and $n = 5$. [27]

According to (14.6), an individual has three different opportunities to manipulate: Telling 0, a value q between $[0, 100]$ or 100. The action in the middle is called "fine-tuning" as the resulting mean coincides with his peak. The new mean reads as follows:

$$\tilde{m}_i(m) = \begin{cases} m - \frac{x_i^*}{5} & m > \frac{6}{5}x_i^* \\ x_i^* & \frac{6}{5}x_i^* - 20 \leq m \leq \frac{6}{5}x_i^* \\ m + \frac{100 - x_i^*}{5} & m < \frac{6}{5}x_i^* - 20 \end{cases} \tag{14.8}$$

The new mean $\tilde{m}_i(m)$ is between m and the peak x_i^*. Whether the payoff with manipulation costs is higher than the payoff under truth-telling depends on the payoff function and is discussed explicitly in what follows.

14.3.4 Mean manipulation with linear payoff functions

In this paragraph, we analyze the best response function and the resulting payoff in case of a linear payoff function. First, we state some general results and in the remainder we consider a concrete example with corresponding figures.

[27] With the adapted values, we have $q := 6x_i^* - 5m$ and the best response

$$BR_{x_i^*}(m) = \begin{cases} 0 & \text{if } u_i\left(m - \frac{x_i^*}{n}\right) - c > u_i(m) \\ q & \text{if } u_i(x_i^*) - c \geq u_i(m) \text{ and } q \in [0, 100] \\ 100 & \text{if } u_i\left(m - \frac{x_i^*}{5} + 20\right) - c > u_i(m) \\ x_i^* & \text{otherwise} \end{cases}$$

Theorem 7. *For a linear payoff function* $u_i(m) = a - b|x_i^* - m|$ *and quasi-fixed costs* $c > 0$ *we have:*
a) If $|x_i^* - m| \leq \frac{c}{b}$*, then* i *has no incentive to manipulate.*
b) If $x_i^* < m - \frac{c}{b}$*, then* i *has no incentive to manipulate if and only if* $x_i^* \leq \frac{nc}{b}$*.*
c) If $x_i^* > m + \frac{c}{b}$*, then* i *has no incentive to manipulate if and only if* $x_i^* \geq 100 - \frac{nc}{b}$*.*

Proof. a) If i manipulates, the maximal utility he can gain is

$$u_i(x_i^*) - c = a - b|x_i^* - x_i^*| - c = a - c \tag{14.9}$$

Hence, if this value is smaller than the utility of telling the truth, he has no incentive to manipulate:

$$a - c \leq a - b|x_i^* - m| \tag{14.10}$$

$$\Leftrightarrow \qquad \frac{c}{b} \geq |x_i^* - m| \tag{14.11}$$

b) Let $x_i^* < m - \frac{c}{b}$. Individual i has no incentive to manipulate if and only if the payoff of the not manipulated mean m is higher than the payoff of the manipulated mean $\tilde{m}$ minus the manipulation costs c.

$$u_i(m) \geq u_i(\tilde{m}_i(m)) - c \tag{14.12}$$

$$a - b|x_i^* - m| \geq a - b|x_i^* - \tilde{m}_i(m)| - c \tag{14.13}$$

As the mean m is larger than the peak x_i^*, individual i proposes a value which is smaller than his peak. Maximal manipulation arises for the smallest value 0. According to Equation (14.8), the manipulated mean is

$$\tilde{m}_i(m) = m - \frac{x_i^*}{n} \tag{14.14}$$

which can be inserted in the utility inequality (14.13)

$$a - b|x_i^* - m| \geq a - b|x_i^* - (m - \frac{x_i^*}{n})| - c \tag{14.15}$$

The constant a arises on both sides and thus can be canceled out. As $x_i^* - m > 0$, the absolute value can be substituted by normal brackets. This leads to

$$x_i^* \leq \frac{nc}{b} \tag{14.16}$$

c) Analogous to part b) with the condition

$$a - b|x_i^* - m| \geq a - b\left|x_i^* - m - \frac{100 - x_i^*}{n}\right| - c \tag{14.17}$$

$$\overset{x_i^* > m + \frac{c}{b}}{\Longrightarrow} \qquad x_i^* \geq 100 - \frac{nc}{b} \tag{14.18}$$

□

Besides knowing on which side of the mean value the own peak is, it is not necessary to know the exact value m. The main statement of the theorem is that the inequalities are independent of m. An individual with an extreme peak, i.e. close to 0 or 100, will manipulate in less cases than an individual with an intermediate peak value.

Example 8 (The optimal payoff). As the best response function is defined via inequalities over the utility function, we start with the visualization of the payoff function, when m is given and individual i with peak x_i^* acts optimal. We consider two different cases: First, we look at the payoff function $u_i(m)$ when i decides not to manipulate and therefore has not to pay any costs. Second, we calculate the payoff when announcing 0, 100 or an optimal value in between minus manipulation costs.

Therefore, we consider the piecewise defined function $u_i(\tilde{m}_i(m)) - c$ (i.e. manipulation with costs) compared with the truth-telling function $u_i(m)$ (without manipulation costs). The special case of $x_i^* = 20$ is plotted in Figure 14.5, more examples can be found in the appendix (see Figure I.1). In all linear examples, let $a = 100$, $b = 1$ and $c = 5$. [28]

The red line shows the payoff when i responds optimal, i.e. $u_i(BR_i(m))$. This function can be intersected in three piecewise linear parts: In the first part, it is optimal to manipulate by "fine-tuning", i.e. by manipulation the new mean coincides with his peak (second line of Equation 14.8. This part is parallel to the $m-$axis and generates a payoff $a - c = 95$. The second part is above the $a - c$ -line. Due to the costs, this area cannot be reached by manipulation (see Part a) of Theorem 7). For $m = x_i^*$ the payoff is maximal and equals $a = 100$. The third part is again the non-manipulation payoff. Here, $x_i^* < m$ and $x_i^* = 20 < \frac{nc}{b} = \frac{5 \cdot 5}{1} = 25$. Hence, we apply Theorem 7, which tells us that i has no incentive to manipulate by telling the minimal value 0. For peaks larger than 25, there exists an interval on the right-hand-side of the peak, such that manipulation to 0 is worth. See the appendix for some illustrating examples. The best response function $BR_{20}(m)$ is plotted in Figure 14.6. It has a jump at $m = 15$. As stated above, for peaks smaller than 25 (Theorem 7.b) or larger than 75 (Theorem 7.c), individuals will never manipulate to the value 0 or 100, respectively.

[28] The figures are simulated and plotted with MatLab.

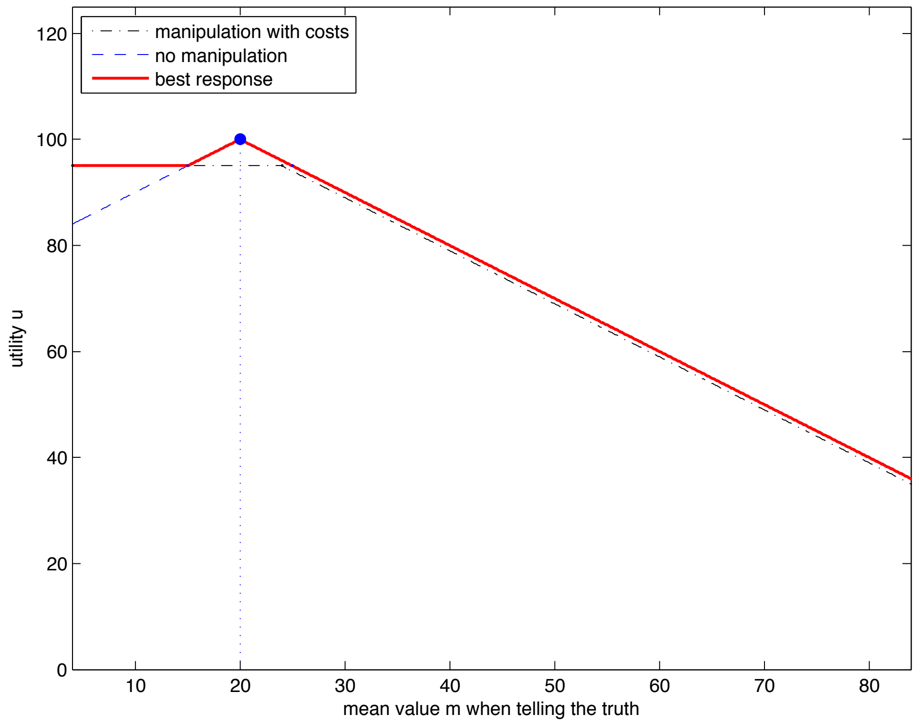


Figure 14.5: Utility of best response to m for $x_i^* = 20$ for a linear payoff function

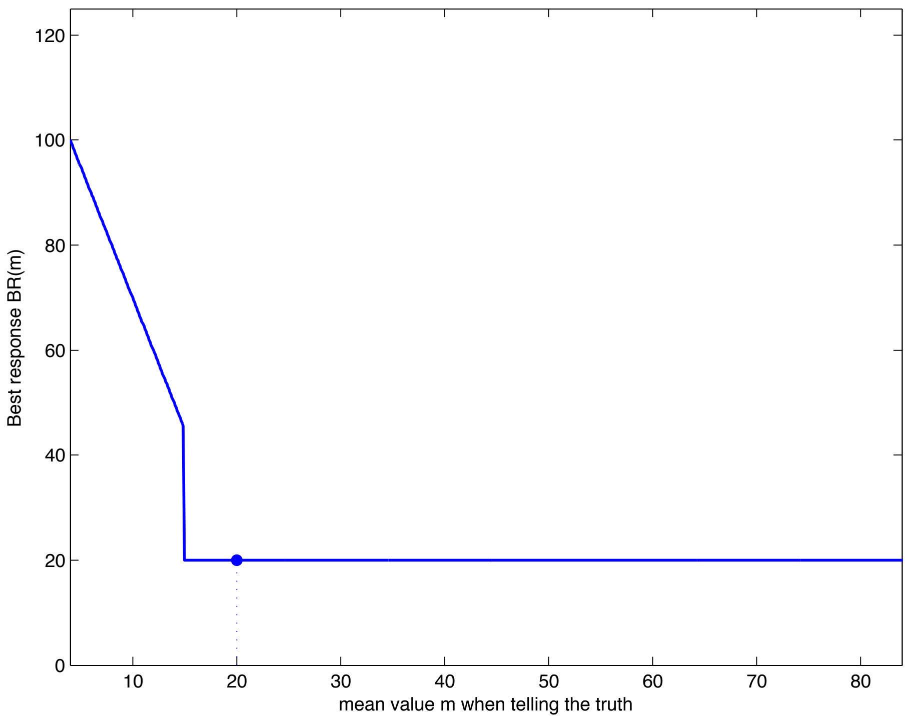


Figure 14.6: Best response function $BR_{20}(m)$ for $x_i^* = 20$ for a linear payoff function

14.3.5 Mean manipulation with quadratic payoff functions

In this paragraph, we consider quadratic payoff functions of the form

$$u_i(m) = a - b(x_i^* - m)^2$$

These functions are differentiable, but still single-peaked. Similar to the linear case, there exists an interval around the peak value, where manipulation is not profitable. The length of this interval is independent of the peak value. Beyond this, truth-telling generates a higher payoff also in a larger interval. This interval depends on the peak value, the parameters of the function and the number of individuals. If the expected mean value is far enough away from the peak value, manipulation is always profitable. A precise description of this vaguely formulated statements is given in the following theorem.

Theorem 8. *For a quadratic payoff function $u_i(m) = a - b(x_i^* - m)^2$ and quasi-fixed costs $c > 0$ it holds:*

a) *If $|x_i^* - m| \leq \sqrt{\frac{c}{b}}$, then i has no incentive to manipulate. In particular, if $|x_i^* - m| > \sqrt{\frac{c}{b}}$ and $\frac{n+1}{n}x_i^* < m < \frac{n+1}{n}x_i^* - \frac{M}{n}$, then the best response is $x_i = x_i^*(n+1) - nm$.*

b) *If $m > \frac{n+1}{n}x_i^*$, then i has no incentive to manipulate if*

$$m \leq \frac{cn}{2bx_i^*} + x_i^*\left(1 + \frac{1}{2n}\right) \tag{14.19}$$

Otherwise he responds optimal with $x_i = 0$.

c) *If $m < \frac{n+1}{n}x_i^* - \frac{M}{n}$, then i has no incentive to manipulate if*

$$m \geq -\frac{cn}{2b(M - x_i^*)} + x_i^*\left(1 + \frac{1}{2n}\right) - \frac{M}{2n} \tag{14.20}$$

Otherwise he responds optimal with $x_i = M$.

Proof. To prove the theorem, we have to compare the utilities of telling the truth with the utility of the best response minus the manipulation costs c. Announcing a value different from either of them, cannot be optimal.

a) The maximal payoff when manipulating equals $u_i(x_i^*) - c = a - c$ and is smaller than the truth-telling payoff $u_i(m) = a - b(x_i^* - m)^2$ if $\sqrt{\frac{c}{b}} \geq |x_i^* - m|$.

b) Let $m > \frac{n+1}{n}x_i^*$, i.e. the maximal manipulation can be achieved by $x_i = 0$, which implies that the new mean equals $m - \frac{x_i^*}{n}$. So truth-telling is better than manipulating if the following condition hold.

$$\begin{aligned} a - b(x_i^* - m)^2 \geq &\quad a - b\left(x_i^* - \left(m - \frac{x_i^*}{n}\right)\right)^2 - c \\ c \geq &\quad -b\left(2(x_i^* - m)\cdot\frac{x_i^*}{n} + \frac{x_i^{*2}}{n^2}\right) \\ m \leq &\quad \frac{cn}{2bx_i^*} + x_i^*\left(1 + \frac{1}{2n}\right) \end{aligned}$$

c) For $x_i^* > m$ the truth-telling condition reads as follows.

$$\begin{aligned} a - b(x_i^* - m)^2 \geq &\quad a - b\left(x_i^* - \left(m + \frac{M - x_i^*}{n}\right)\right)^2 - c \\ \frac{cn}{b(M - x_i^*)} \geq &\quad -\left(-2x_i^* + 2m + \frac{M - x_i^*}{n}\right) \\ m \geq &\quad -\frac{cn}{2b(M - x_i^*)} + x_i^*\left(1 + \frac{1}{2n}\right) - \frac{M}{2n} \end{aligned}$$

□

Example 9. For $a = 200$, $b = 0.02$ and $c = 5$, truth-telling is a best response for all x_i^* whenever m is within $[x_i^* - 5\sqrt{10}, x_i^* + 5\sqrt{10}]$.
Figure 14.7 and 14.8 illustrate the example of $x_i^* = 20$, more figures can be found in the appendix. In Figure 14.8, we see that for this choice of parameters, there is no fine-tuning interval: only the values $\{0, x_i^*, 100\}$ are optimal, respectively. This is a peculiarity of this concave payoff function. Let us consider values of $m > \frac{n+1}{n}x_i^* = 1.2 \cdot 20 = 24$. Without costs, it would be optimal to announce the minimal value 0. By Equation 14.19, truth-telling for large m is better, whenever

$$m \leq \frac{25}{0.04 \cdot 20} + 20 \cdot \frac{11}{10} = 53.25$$

This is exactly the value, for which the utility of best response switches from "no manipulation" to "manipulation with costs" in Figure 14.7.

As in the linear case, the best response function is a decreasing function in m. However, this is not true for every payoff function as we will see in the following paragraph.

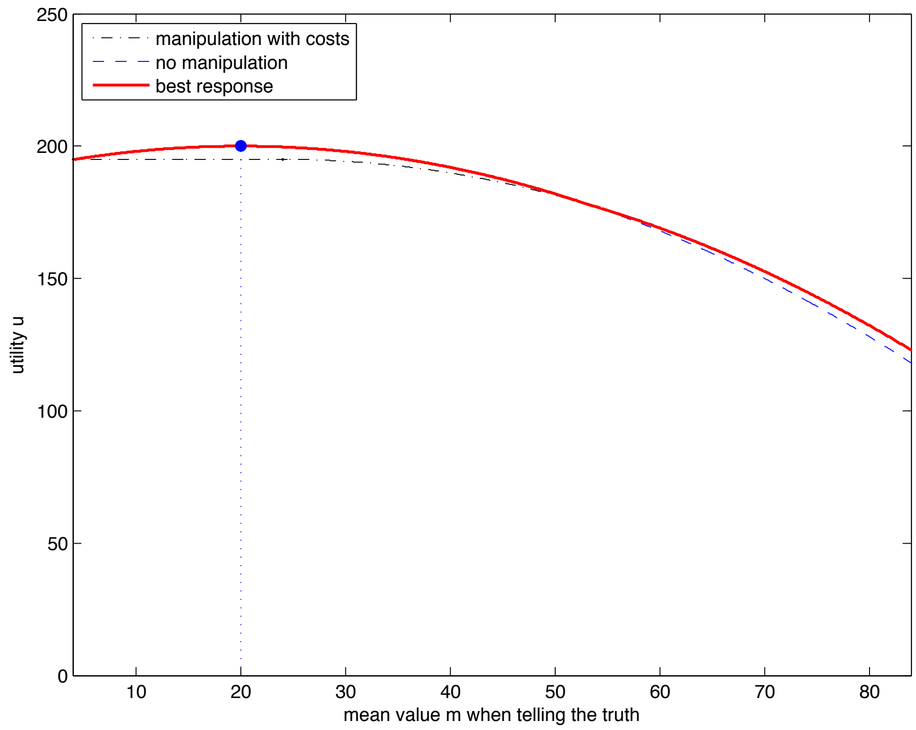


Figure 14.7: Utility of best response to m for $x_i^* = 20$ for a quadratic payoff function

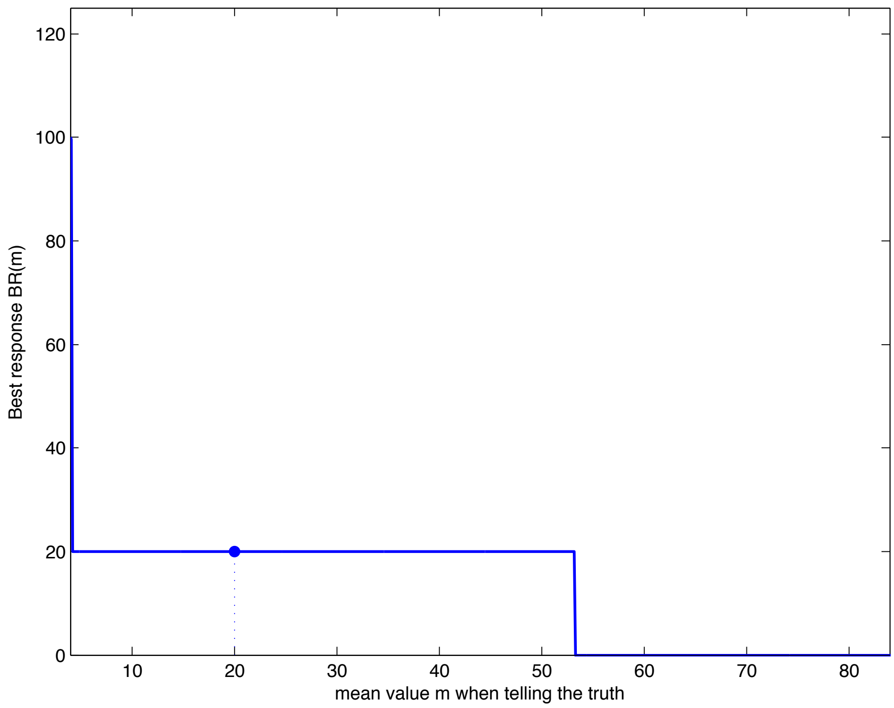


Figure 14.8: Best response function $BR_{20}(m)$ for $x_i^* = 20$ for a quadratic payoff function

14.3.6 Mean manipulation with special payoff functions

As a last example, we consider the special payoff function. We start with a theoretical result and explain its meaning in a concrete example afterwards.

Theorem 9. *For the special payoff function*

$$u_i(m) = 10 + \min\left(\frac{380}{|m-(x_i^*-2)|}, \frac{380}{|m-(x_i^*+2)|}\right) \tag{14.21}$$

and quasi-fixed costs $c > 0$ it holds:

a) If $|x_i^* - m| \le \frac{2}{37}$, then i has no incentive to manipulate. In particular, if $|x_i^* - m| > \frac{2}{37}$ and $\frac{n+1}{n}x_i^* < m < \frac{n+1}{n}x_i^* - \frac{M}{n}$, then the best response is $x_i = x_i^*(n+1) - nm$.

b) If $m > \frac{n+1}{n}x_i^*$, then i has no incentive to manipulate if

$$m \ge -2 + x_i^*\left(1 + \frac{1}{2n}\right) + \sqrt{\frac{x_i^*}{n}\left(\frac{x_i^*}{4n} + \frac{380}{c}\right)} \tag{14.22}$$

Otherwise he responds optimal with $x_i = 0$.

c) If $m < \frac{n+1}{n}x_i^* - \frac{M}{n}$, then i has no incentive to manipulate if

$$m \le 2 + x_i^*\left(1 + \frac{1}{2n}\right) - \frac{50}{n} - \sqrt{\frac{100 - x_i^*}{n}\left(\frac{100 - x_i^*}{4n} + \frac{380}{c}\right)} \tag{14.23}$$

Otherwise he responds optimal with $x_i = M$.

Proof. a) Maximal payoff is obtained when the manipulated mean $\tilde{m}_i(m)$ equals the peak.

$$u_i(x_i^*) - c = 10 + \frac{380}{x_i^* - x_i^* + 2} - 5 = 195$$

By comparing this value with truth-telling, we obtain

$$\begin{aligned}
& 10 + \min\left(\tfrac{380}{|m-(x_i^*-2)|}, \tfrac{380}{|m-(x_i^*+2)|}\right) \ge 195 \\
\Rightarrow \quad & \min\left(\tfrac{1}{|m-x_i^*+2|}, \tfrac{1}{|m-x_i^*-2)|}\right) \ge \tfrac{185}{380} \\
\Rightarrow \quad & \tfrac{380}{185} \ge \max(|m - x_i^* + 2|, |m - x_i^* - 2)|) \\
\Rightarrow \quad & \tfrac{76}{37} \ge |m - x_i^*| + 2 \\
\Rightarrow \quad & |m - x_i^*| \le \tfrac{2}{37}
\end{aligned}$$

b) If $m > \frac{n+1}{n} x_i^*$, the relevant part of the min-function is the first argument. Truth-telling is better than manipulating if

$$\frac{380}{|m - x_i^* + 2|} \geq \frac{380}{\left|m - \frac{x_i^*}{n} - x_i^* + 2\right|} - c$$

$$\Rightarrow \quad m - \frac{x_i^*}{n} - x_i^* + 2 \geq m - x_i^* + 2 - \frac{c}{380}(m - x_i^* + 2)\left(m - \frac{x_i^*}{n} - x_i^* + 2\right)$$

$$\Rightarrow \quad m^2 + m\left(4 - 2x_i^* - \frac{x_i^*}{n}\right) + \left((2 - x_i^*)^2 + \frac{x_i^{*2}}{n} - \frac{2x_i^*}{n} - \frac{380x_i^*}{cn}\right) \geq 0$$

This inequality has two values for which it holds with equality. Due to the structure of our problem, we are interested in the larger one.

$$m \geq -2 + x_i^* + \frac{x_i^*}{2n} + \sqrt{\left(2 - x_i^* - \frac{x_i}{2n}\right)^2 - \left((2 - x_i^*)^2 + \frac{x_i^{*2}}{n} - \frac{2x_i^*}{n} - \frac{380x_i^*}{cn}\right)}$$

$$\Rightarrow \quad m \geq -2 + x_i^*(1 + \frac{1}{2n}) + \sqrt{\frac{x_i^*}{n}\left(\frac{x_i^*}{4n} + \frac{380}{c}\right)}$$

c) analogous. □

Example 10. Figure 14.10 illustrates the best response function and Figure 14.9 the utility of the best response function for $x_i^* = 20$.

The special payoff function reacts very sensitive on manipulation, i.e. in a large interval around the peak, manipulation generates a higher payoff than truth-telling. A mean value far away from the peak may imply truth-telling as the generated benefit does not compensate manipulation costs. In the direct neighborhood of x_i^*, there exists a small interval, where truth-telling is the best response. To be precise, this occurs when $|m - x_i^*| \leq \frac{2}{37}$. This peculiarity can be seen in utility function (Figure 14.9) but is too small to be noticed in the best response function (Figure 14.10) and is also too small for being relevant in the experiments.

We also notice that the fine-tuning interval, i.e. the plateau at the utility level $u_{max} - c = 195$ coincides relatively often with the best response utility. As an implication, the best response of one individual may have counter-productive effects on another individual.

In this example (Figure 14.10) the marginal utility of the special payoff function for large values is very small as the payoff function is piecewise convex. Hence, manipulation is not profitable, and we observe again a discontinuity.

14.4 Comparing the two rules

We have seen that mean and median rules generate social outcomes which are "somewhere in between" of the suggested values. They are somewhat similar since the

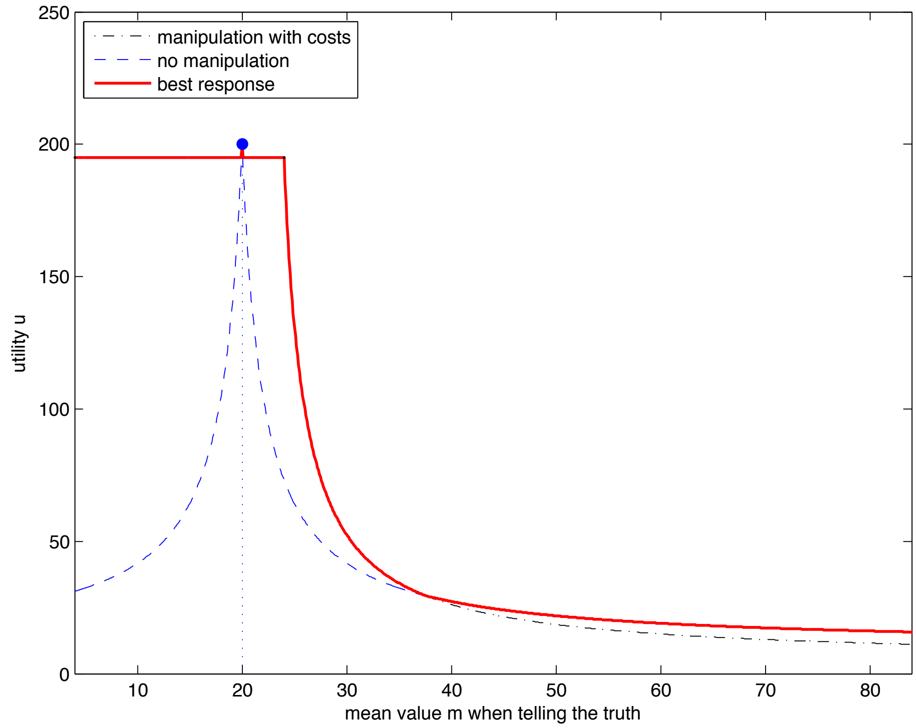


Figure 14.9: Utility of best response to m for $x_i^* = 20$ for the special payoff function

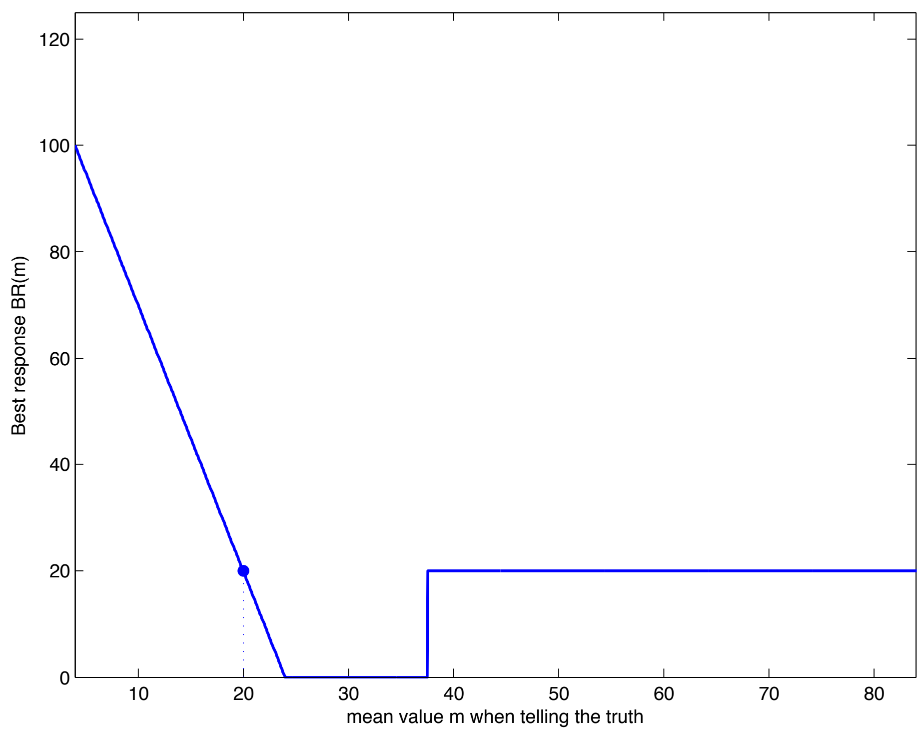


Figure 14.10: Best response function $BR_{20}(m)$ for $x_i^* = 20$ for the special payoff function

median can be written as the $\frac{n-1}{2n}$-truncated mean. However, the dissimilarities of both rules are more conspicuous. The mean rule is manipulable whereas the median rule is not. Without costs, there exists a unique and efficient Nash equilibrium under the mean rule, while inefficient equilibria are also possible under the median rule. As the median rule is strategy-proof, it is not necessary to have full information over the peaks of the other individuals to state an optimal proposal. However, under the mean rule, individuals have to know at least their ranking number within all peaks to calculate their optimal choice. This includes knowledge of the total number of individuals n. Moreover, the assumption of the rational behavior of other individuals plays a crucial role in calculating the equilibrium. Also, the outcome value differs: Under the median rule, it corresponds to one of the proposed values, whereas under the mean rule it may be a different value. An overview over these properties is given in Table 14.2 where both rules are compared with and without costs.

	mean		median	
	no costs	costs	no costs	costs
strategy-proof	No	No	Yes	Yes
equilibrium always exists	Yes	No	Yes	Yes
if exists, unique	Yes	No	No	Yes
inefficient equilibria possible	No	Yes	Yes	No
more peak info than own peak necessary	Yes	Yes	No	No
more payoff function info than own peak necessary	No	Yes	No	No

Table 14.2: Mean and Median rule compared

We briefly highlight some of the results shown in Table 14.2. Depending on the distribution of peaks, there may be no equilibrium in pure strategies for the mean rule with costs. This exception arises, when there are at least a pair of agents of those one manipulates if and only if the other does not and vice versa. Such a cycle occurs for instance in the following example.

Example 11. Table 14.3 shows an illustrative example, where it depends on the payoff function whether a unique equilibrium exists or not.[29] The peaks of the five individuals are $10, 20, 50, 60, 70$. Hence, for every payoff function without manipulation costs, there exists a unique equilibrium with mean 50 (yellow row). The table

[29]If two peaks are very close together, it may happen that the same mean value occurs when telling the truth or manipulating, but unilateral deviation is not profitable. This is a very special case and not relevant for our analysis.

Individual	1	2	3	4	5	Mean
Peak	10	20	50	60	70	42
ewoc	0	0	50	100	100	50
	Linear payoff function					
$m_i(50)$	52	54	50	42	44	
$BR_{x_i^*}(m_i)$	10	20	50	100	100	56
$m_i(56)$	56	56	56	48	50	
$BR_{x_i^*}(m_i)$	10	20	20	100	100	50
$m_i(50)$	50	50	56	42	44	
equilibrium	10	20	20	100	100	50
	Quadratic payoff function					
$m_i(50)$	52	54	50	42	44	
$BR_{x_i^*}(m_i)$	10	0	50	60	100	44
$m_i(44)$	44	48	44	44	38	
$BR_{x_i^*}(m_i)$	10	20	74	60	100	52.8
$m_i(52.8)$	52.8	52.8	48	52.8	46.8	
$BR_{x_i^*}(m_i)$	10	20	50	60	70	42
$m_i(42)$	42	42	42	42	42	
$BR_{x_i^*}(m_i)$	10	20	50	60	100	48
$m_i(48)$	48	48	48	48	42	
equilibrium	10	20	50	60	100	48
	Special payoff function					
$m_i(50)$	52	54	50	42	44	
$BR_{x_i^*}(m_i)$	10	20	50	100	70	50
$m_i(50)$	50	50	50	42	50	
$BR_{x_i^*}(m_i)$	10	20	50	100	100	56
$m_i(56)$	56	56	56	48	50	
$BR_{x_i^*}(m_i)$	10	20	20	100	100	50
$m_i(50)$	50	50	56	42	44	
$BR_{x_i^*}(m_i)$	10	20	20	100	70	44
$m_i(44)$	44	44	50	36	44	
circle	10	20	50	100	70	50

Table 14.3: Example of non-existing equilibria

shows the iterative calculation for the three different payoff functions, we analyzed in the previous paragraphs. As initial value we use the constellation of the equilibrium without costs (*ewoc*). As in the sections before, for each person, the best response function $BR_{x_i^*}$ is based on the mean value if he announced the truth. This value is calculated by taking the actual mean, subtracting the announced value and adding the peak value in appropriate proportions. Hence, it is determined by

$$m_i(m) := m + \frac{x_i^*}{5} - \frac{x_i}{5}$$

where x_i is the last calculated best response or the initial value in the first round. We need this intermediate step to apply the results of the previous sections. For instance, individual 1 announces 0 in the *ewoc*, i.e. if he told the truth, the mean would be 52. In the *ewoc*, only individual 3 tells the truth, so for he the mean value (50) does not change.

In the next row, the best response is given according to our previous analysis for the linear payoff function. We see that individuals 1 and 2 no longer have an incentive to manipulate and hence tell the truth, whereas it is profitable for individual 4 and 5 still to deviate to the maximal value 100. This steps are repeated until we see that the best response function does not change and the stable equilibrium $(10, 20, 20, 100, 100)$ with mean 50 is reached.

For the quadratic payoff function, the general idea remains the same, but more iterations are necessary and a different equilibrium $(10, 20, 50, 60, 100)$ with mean 48 results (green rows).

In the lower part of the table, it is shown that for the special payoff function a circle occurs. The last line coincides with the initial best response values (red rows). Every time, individual 3 tells the truth, individual 5 has an incentive to deviate and vice versa. The exact proof is given in the Appendix H, here only a sketch of the proof is given. In an equilibrium it must hold that $BR_{10} = 10$, $BR_{20} = 20$ and $BR_{60} = 100$ which can be shown by contradiction. Hence, only the choices of individuals 3 and 5 are flexible and here no equilibrium is possible as explained above. We conclude that there exist equilibria for the linear and quadratic payoff function, but not for the special one.

As we saw in Section 14.3.3, the calculation of the best response function can be complex for the mean rule with costs. For each of the studied payoff functions, it was necessary to have information about the expected mean value. This implicitly signifies a knowledge about the others' peaks. Manipulating or not and therefore the best response function itself is sensitive to the structure and parameters of the payoff function. This fact must be taken into account when updating the belief over the expected mean value. On the other hand, the median rule is very robust in all those mentioned issues. Combining individuals with different payoff parameters or even different types of payoff functions has no influence on the resulting equilibrium. Although complexity is not a topic of this work, obviously the calculation of the median equilibrium is very easy and can become very difficult for mean equilibria with costs. Attention should be paid to this when analyzing the experimental results.

As stated in Table 14.2 there exist inefficient equilibria for the mean rule with manipulation costs and for the median rule without costs. The latter one has been already discussed in Section 14.2.1, for the mean rule we give the following example.

Example 12. Assume that the peaks are given by $(10, 20, 45, 60, 70)$. Then, the mean rule with the special payoff function and manipulation costs of $c = 5$ leads to the equilibrium $(10, 20, 25, 100, 70)$ (see also Table G.1). But $z = (10, 20, 65, 60, 70)$ would not change the mean value (45) but increase the payoff of the individual 4 with peak 60, as he has not to pay manipulation costs. Note that z is not an equilibrium as individual 4 has an incentive to manipulate. Hence, this equilibrium is inefficient.

15 Welfare effects and maximal payoffs under different payoff functions

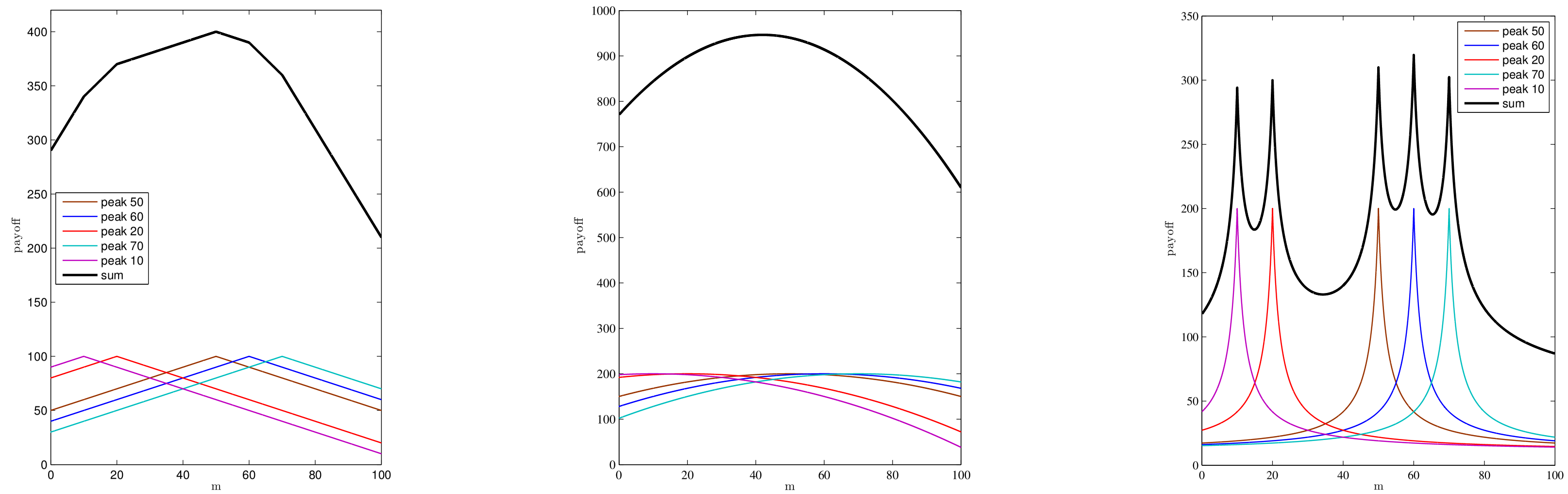
As a small excursion, we analyze in this section the overall payoff if all individuals have payoff functions of the same form $u_i(m)$ with different peaks $x_1^*, \ldots, x_n^*$. Depending on the form of individual payoffs the sum, i.e. the utilitarian social welfare functional $U(m) = \sum_{i=1}^{n} u_i(m)$, also has distinct properties. We consider linear, quadratic and special payoff functions as introduced in Equations (14.1)-(14.3). For this analysis, it is not relevant whether m is the mean, median or any other value and therefore we do not apply manipulation costs.

15.1 Linear payoff functions

First, we consider the case of piecewise linear functions, i.e. functions of the form $u_i(m) = a - b|x_i^* - m|$ with fixed values a and b and individual peaks x_i^*. Because of concavity, the sum of the functions $U(m) = \sum_{i=1}^{n} u_i$ is also piecewise linear and single-peaked. Its maximum is at the median of the individual peaks. Figure 15.1(a) shows an example for $a = 100$, $b = 1$ and the peaks $(10, 20, 50, 60, 70)$. The maximum of the sum is at the median of the peak values $m = 50$.

15.2 Quadratic payoff functions

Second, we analyze quadratic payoff functions, i.e. utility functions which can be written as $u_i(m) = a - b(x_i^* - m)^2$. The sum $U(m) = \sum_{i=1}^{n} u_i(m)$ is also quadratic as we can easily see:



(a) Linear payoff functions and their sum: Maximum at the median

(b) Quadratic payoff functions and their sum: Maximum at the mean

(c) Payoff functions used in the experiment and their sum: Maximum at one peak

Figure 15.1: Payoff functions and their sum for peak values $(10, 20, 50, 60, 70)$

$$
\begin{aligned}
U(m) =& \sum_{i=1}^{n} u_i(m) \\
=& \sum_{i=1}^{n} a - b(x_i^* - m)^2 \\
=& na - b \left(\sum_{i=1}^{n} x_i^{*2} + 2m \sum_{i=1}^{n} x_i^* - nm^2 \right)
\end{aligned}
$$

For maximizing the sum, we need the set the first derivation of U equal to zero.

$$
\begin{aligned}
\frac{\partial U}{\partial m} = U'(m) = -2b \sum_{i=1}^{n} x_i^* + 2bnm = 0 \\
m = \frac{1}{n} \sum_{i=1}^{n} x_i^* = \bar{x}^*
\end{aligned}
$$

This signifies that overall payoff is maximal for the mean of the individual peaks. Figure 15.1(b) shows an example with the same peaks as in the linear case. We adapted the parameters to $a = 200$ and $b = 0.02$. The maximum of the sum function is at the mean value $m = 42$.

15.3 Special payoff functions

The last type of payoff functions we analyze here, is the one we used in the experiments. This function is of the form

$$
u_i(m) = 10 + \min \left(\frac{380}{|m - (x_i^* - 2)|}, \frac{380}{|m - (x_i^* + 2)|} \right)
$$

The function $u_i(m)$ is piecewise hyperbolic and symmetric with respect to $x = x_i^*$. In contrast to the previous examples, the sum $U(m) = \sum_{i=1}^{n} u_i(m)$ is not single-peaked any more. One of the reasons is that the functions $u_i(m)$ are not concave. The maximal value of $U(m)$ is reached at one of the peaks and depends on the distribution of the peaks and the distances between them. Figure 15.1(c) shows the special payoff function for the peaks $(10, 20, 50, 60, 70)$ as used in the previous examples. Here, the maximum of the sum function is achieved for $m = 60$. Because of the piecewise monotonicity of the u_i's, the value maximizing $U(m)$ can never be one of the extreme peaks x_1^* or x_n^*. Hence, in the case of $n = 3$, the maximum coincides with the median.[30]

[30] A comprehensive analysis does not really lead to "nice" equations but to long formulas without added value.

16 Experimental Design

In order to test whether the theoretical results hold also in practice, we conducted an experimental study.[31] In this section we first explain the general setup of the experiment and the outline of the experimental procedure. We then explain the different treatments. Our main treatment variables are "rule", "info" and "cost". We also have secondary treatment variables such as "experiment" and "framing". The first is an auxiliary variable, the second was only used in the *no cost* treatments. They are explained in detail in the following subsections.

16.1 General setup

All sessions took place at Karlsruhe Institute of Technology in 2012. Each experiment consisted of a "Mean" and a "Median" part. The students were divided into groups of five. Two groups were in the laboratory at the same time and we ensured anonymous interaction.

In total we observed the behavior of 235 participants (see Table 16.1 for details). Students, who showed up, but did not participate received a show-up fee of 5 €. The average wage of participants was 14.26 € for the duration of approximately one hour.

To implement single-peaked preferences, all participants were paid according to the same payoff function (see Equation 14.3) which depended on the group decision and their individual peak. The function was introduced to the students at the beginning of the experiment and denominated in "*ECU*"[32].

Given that proposals between 0 and 100 were allowed, feasible payoffs were between 10 *ECU* and 200 *ECU*.

[31] We thank the "Fondation Université de Strasbourg" for the financial support.

[32] Experimental Currency Unit

16.2 Laboratory procedure

We recruited the participants via ORSEE (see Greiner (2004)). The experiments took place in the experimental laboratory of the department of economics at KIT. The workplaces were allocated arbitrarily to the participants after arrival. Separating walls prevented the persons from talking to each other. Figure 16.1 shows a typical workplace equipped with a PC, paper and pencil. The calculator was integrated in the zTree (see Fischbacher (2007)) environment and available all the time.

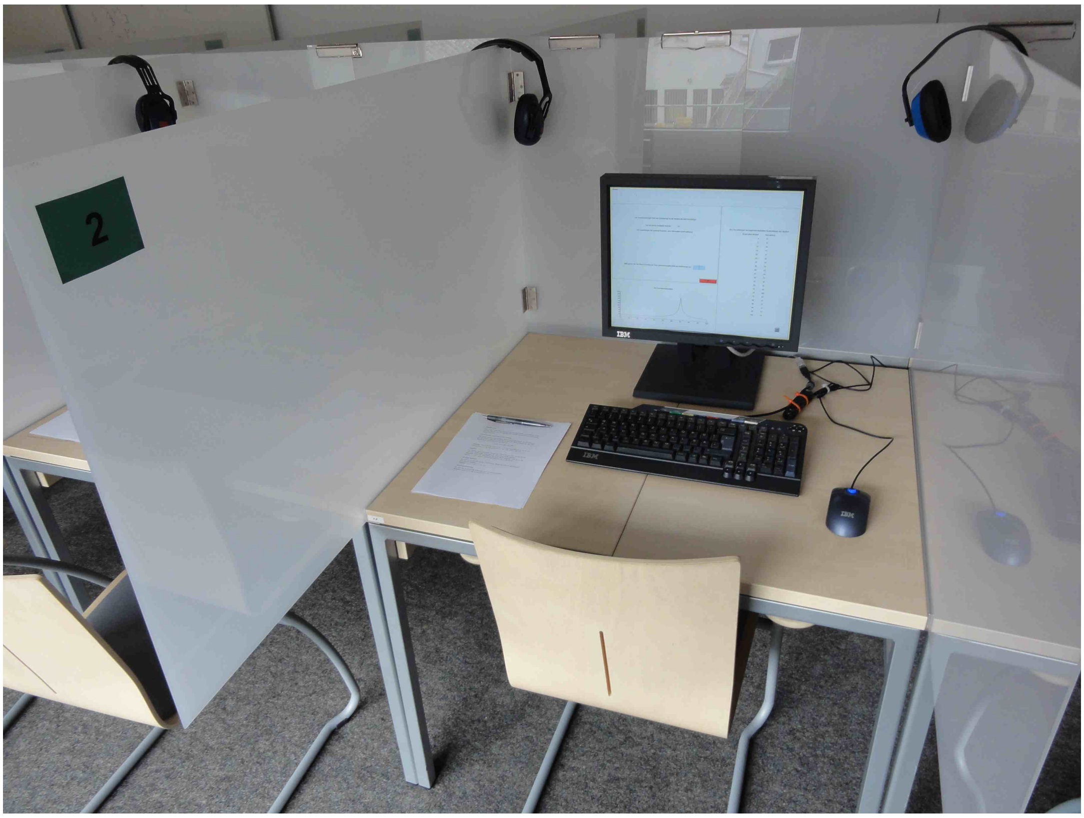


Figure 16.1: Workplace at the KIT laboratory

First, the mathematical instructions (see Appendix K) were read out aloud to the participants. We used an audiotape to guarantee constant quality and to avoid any variation across sessions. Then, participants answered test questions on the PC to assure that they understand the calculation of mean and median values and basic properties of their payoff function. After solving all tasks correctly, the instructions (see Appendix L) were read aloud. The instructions were different depending on the framing (see Section 16.4.5). During the entire experiment, the participants were able to read the instructions again on the handout. Then, each participants had to make 48 decisions[33] according to the peak distribution explained in Section 16.4.6. At the end of the experiment, they filled in a short questionnaire with demographic characteristics. The students were payed one by one, such that they could not observe the profit of the others.

16.3 Demographic data

In total, we observed the behavior of 235 participants of whom 67 were female. This is representative for the Karlsruhe Institute of Technology where there are about 27% female students. Out of all participants, 160 ticked that their subject of study is related to economics, i.e. either major or minor subject. Minimal age was 17, maximum age 37, so we got an average of 22.4 years. At the end of the study, the students were asked whether it was possible to manipulate under the mean rule (221 correctly answered “yes”) and whether it was possible to manipulate under the median rule (168 correctly answered “no”).

16.4 Treatment variables

16.4.1 Treatment variable “rule”

The treatment variable **rule** is a within-subject treatment variable. It determines which rule is used to aggregate the individual values to a group value. It can take the values *mean* and *median*. Both rules were explained in the previous sections.

16.4.2 Treatment variable “info”

The treatment variable **info** is a within-subject treatment variable. *Info* is used as the abbreviation for *information* and determines whether information about the peak of the others is available or not. It can take the values *no info* and *full info*. In rounds with *no info* (NI), participants knew their own peak but not the peak of the other participants. In rounds with *full info* (FI), participants in addition knew the peaks of the other four participants. The type of information was the same for all participants, i.e. either every one had no information or full information for a single decision. As we already explained in Section 14, full information is necessary to find the mean equilibrium but not required for the median equilibrium.

[33] Data were collected with zTree software. Due to a mistake, two peak distributions were replayed in the first experiment. Treatment 8 (session 8) and 10 (session 5) were played twice. The corresponding data were deleted for the analysis.

16.4.3 Treatment variable "cost"

The treatment variable **cost** is a between-subject treatment variable. It can take the values *no costs* and *costs* and indicates whether manipulation costs occured. In experiment 1 and 2 there were no manipulation costs. In experiment 3 and 4 we introduced costs of 5 *ECU* per decision if the proposed value was unequal the peak. So we added this sentence to the introduction of the abstract framing:[34]

If your proposal x is different from your value x_i^*,
you incur expenses of 5 *ECU in this round.*

Changes in the Nash equilibria were already discussed in Sections 14.2.2 and 14.3.3. An overview of the equilibria is given in Table G.1.

16.4.4 Treatment variable "exp"

The treatment variable **exp** is a between-subject auxiliary treatment variable. It is used to distinguish between the sequence of the treatment **rule** in the different treatments of **cost** and to check for sequence effects.

- exp.1: *no costs*, where **rule** is played in the sequence *mean-median*.
- exp.2: *no costs*, where **rule** is played in the sequence *median-mean*.
- exp.3: *costs*, where **rule** is played in the sequence *mean-median*.
- exp.4: *costs*, where **rule** is played in the sequence *median-mean*.

Unless otherwise stated, (exp.1 and exp.2) are grouped in *no costs* and (exp.3 and exp.4) are grouped in *costs* in the following analysis.

16.4.5 Treatment variable "framing"

The treatment variable **framing** is a between-subject treatment variable. It can attain four different values: *A, F, J, Z*. In each session, the task was explained to the participants in one particular framing.

- Abstract (A): The first framing was the control group. The instructions were neutral, in the sense that we used the terms "value", "payoff" and so on. This was the only framing which was used for the treatment "costs".

[34]original German version: "Wenn Ihr Vorschlag y von Ihrem Wert w verschieden ist, entstehen Ihnen in dieser Runde Kosten in Höhe von 5 *ECU*." In the experiments, we used "y" instead of "x" and "w" instead of "x_i^*" to avoid confusion. In this work, we try to label the variables consistent with the denomination introduced in the theoretical part.

		framing				
cost	**exp**	*Abstract*	*Company*	*Jury*	*Bank*	sum
no costs	1	30	20	25	20	95
	2	20	30	20	30	100
costs	3	20	0	0	0	20
	4	20	0	0	0	20
sum		**90**	**50**	**45**	**50**	**235**

Table 16.1: Number of participants per framing

- Company (F): In treatment F, a "manager" had to decide where to construct a "shop" for selling the products manufactured at the headquarters' position. His profit depended on the distance of the headquarters to the shop because of "transportation costs".
- Jury (J): In treatment J, a "jury member" had to score "athletes" according to their performance.
- Bank (Z): In treatment Z, an "expert" of the central bank had to estimate the seasonally adjusted change of the money supply compared to the previous year.

Depending on the framing, participants received their paper instructions in a different wording. The terms used on the computer monitor were adapted as well.[35] Table 16.1 shows the number of participants for each framing in all three experiments. In each treatment, at least 20 persons participated.

16.4.6 Treatment variable "distribution of peaks"

The distribution of peaks is another auxiliary variable which is shown in Table 16.2. As peaks were not assigned randomly, it is possible to look at each peak distribution separately and compare the results between individuals. There exist basically four different distributions but as only one peak was assigned to each player per round, we obtain 12 distribution numbers by permutation. For instance, distribution number 1 and 5 both consist of the peaks $(10, 20, 50, 60, 70)$ but no player had the same peak in these distributions (player 1: 50 and 70, player 2: 60 and 10 etc.) The *no info* treatments were repeated five times each, the *full info* treatments three times. The players did not know their player number and were informed about their peak not until the corresponding period. The number of repetitions was not announced in advance. In the right column, the corresponding decision numbers are listed. For each player, a list of 48 decisions is obtained.

[35] A detailed instruction in German for the framing "abstract" is attached in the Appendix L . All other instructions are available upon request.

			player					
Distr. no.	treatment		1	2	3	4	5	Dec.no.
1	*mean*	*no info*	50	60	20	70	10	1-5
2			70	60	50	30	65	6-10
3			10	20	30	40	50	11-15
4			60	70	10	20	45	16-20
5		*full info*	70	10	60	50	20	21-23
6			30	50	70	65	60	24-26
7			40	50	10	20	30	27-29
8			20	10	45	60	70	30-32
9	*median*	*no info*	20	70	50	10	60	33-37
10			60	45	20	70	10	38-42
11		*full info*	30	40	50	10	20	43-45
12			70	50	65	60	30	46-48

Table 16.2: Distribution of peaks (Distr.no.) and corresponding decision numbers (Dec.no)

17 Research questions

Before we summarize our data in the next chapter, we shortly state our main hypotheses.

1. **Truth-telling under the median rule.** According to the theoretical analysis, truth-telling is a weakly dominant strategy and should therefore be played often under the median rule. It is not possible to verify this in a t-test or a weaker sign rank test because of the structure of the question: If we compare the average deviation, i.e. the value of the peak minus the suggested value, with the normal distribution with expected value 0, strategically motivated deviations cannot be distinguished from stochastic deviations. Therefore, we give a more descriptive analysis of this hypothesis.
2. **There is more truth-telling under the median rule than under the mean rule.** Compared to the previous hypothesis, the negation of this hypothesis can be rejected by an appropriate hypothesis test. We apply the t-test for paired samples to compare the average absolute deviation per group under the median and the mean rule restricted to information and costs.
3. **Unique Nash equilibrium under the mean rule.** As explained in Section 14.3.1, we assume that when there are no manipulation costs, the unique Nash equilibrium is played under the mean rule. We count how often individuals announce these values and analyze the results in a descriptive way.
4. **Costs increase truth-telling under the mean rule and under the median rule under full information.** Cost reduce average absolute deviation per group. We apply a two sample Wilcoxon rank-sum (Mann-Whitney) test.
5. **Costs reduce Nash play under the mean rule.** It is more challenging to determine the equilibrium strategy when there are costs. We apply a two sample Wilcoxon rank-sum (Mann-Whitney) test.
6. **There are no sequencing effects between median-mean and mean-median.** We run the Mann-Whitney-U test for two independent samples to check for sequencing effects in the average absolute deviation from the peak.
7. **Full information reduces truth-telling under the mean rule.** In a within-subject comparison, we check the effect of information on absolute deviation of the peak by t-tests.

8. **Framing has no influence.** In our regression model, we use framing as an independent variable to show that it does not influence the absolute deviation from the peak.

18 Summary of our data

In the previous sections we explained the theory of two important voting rules and introduced the design of our experiment. The goal of this section is the presentation of the results in a descriptive way. We classify the announced values of the participants in different “strategy-types”, where strategy is a term we use for the one-shot game. When using the term “Nash play”, we refer to the equilibrium action as determined in the theoretical part of this work.

This section consists of three parts: In the first part, we analyze the aggregated group outcome to figure out whether the aggregated mean and median values coincide with the predicted values. In the second part, we look at the individual actions and aggregate these results independently of the person announcing this value. This is the main part of the analysis as we are able to make general statements on truth-telling, Nash play and other strategies. Finally, in the third part, we take a closer look at the individuals themselves. *Do some individuals tell the truth more often than others?* is one of the questions we consider in that part.

18.1 Actual aggregated group outcome

18.1.1 Achieving the equilibrium outcome

Independently of the individual actions, it is an interesting issue, whether the social outcome predicted by the equilibrium analysis is achieved in the experiments. Table 18.1 presents the percentage of coincidence for all different treatments. Values in brackets show coincidence with an approximated value which is in the interval of the equilibrium value ± 5.

In the mean treatments, the underlying theoretically predicted Nash equilibrium outcome is the same for experiments with and without manipulation costs. Over all experiments, the mean value hardly equals the predicted Nash value ($3\% - 8\%$). These values are significantly higher, when the approximated values are taken into account ($40\% - 50\%$).

The exact median of peaks was achieved more often ($49\% - 75\%$). An approximate value was obtained in $64\% - 100\%$ of median decisions. These values are plotted

		mean		*median*	
		no info	*full info*	*no info*	*full info*
no costs	exp. 1	0.04 (0.47)	0.04 (0.45)	0.53 (0.64)	0.49 (0.75)
	exp. 2	0.06 (0.50)	0.05 (0.40)	0.66 (0.80)	0.68 (0.91)
costs	exp. 3	0.03 (0.50)	0.08 (0.50)	0.63 (0.68)	0.50 (0.67)
	exp. 4	0.05 (0.48)	0.08 (0.44)	0.60 (0.65)	0.75 (1.00)
	all	0.05 (0.49)	0.05 (0.43)	0.60 (0.71)	0.60 (0.83)

Table 18.1: Percentage of decisions, in which the mean/median of the announced values equals the Nash equilibrium (or approximated values)

over time in Figure 18.1 and 18.2 when exp.1 and exp.2 are aggregated (no costs) and exp.3 and exp.4 are aggregated (costs). On the horizontal axis, the decision numbers over time are plotted (see Table 16.2). The numbers refer to the first round of a peak distribution. Therefore, under no information every fifth round and under full information every third round is labeled. As only group decisions were counted, the underlying datasets consist of 39 values without costs and 4 values with manipulation costs. Therefore, the results of *costs* should be used for comparison only.

18.1.2 Average aggregated values

Figure 18.3 shows the actual outcome and the outcome in the equilibrium situation (green) over all rounds. The blue circles represent the treatments without costs, the red triangles those treatments with manipulation costs. Changes in the peak distribution can be seen at the steps of the green squared function and again by the labeling of the abscissa.

Compared to the previous section in which we counted how many times the equilibrium value was achieved, we now illustrate the average actual group values. A special value is decision number 11, where the average mean value is conspicuously smaller than the equilibrium value. One possible explanation is the positive skew of the peak distribution: Participants may belief that their peak is smaller than the peak of others and therefore manipulate in the wrong direction. In decision 27, there is the same peak distribution but full information and we observe that the difference between actual outcome and equilibrium value is very small in the treatment

Figure 18.1: Percentage of Mean decisions, in which the mean of the announced values equals the Nash equilibrium

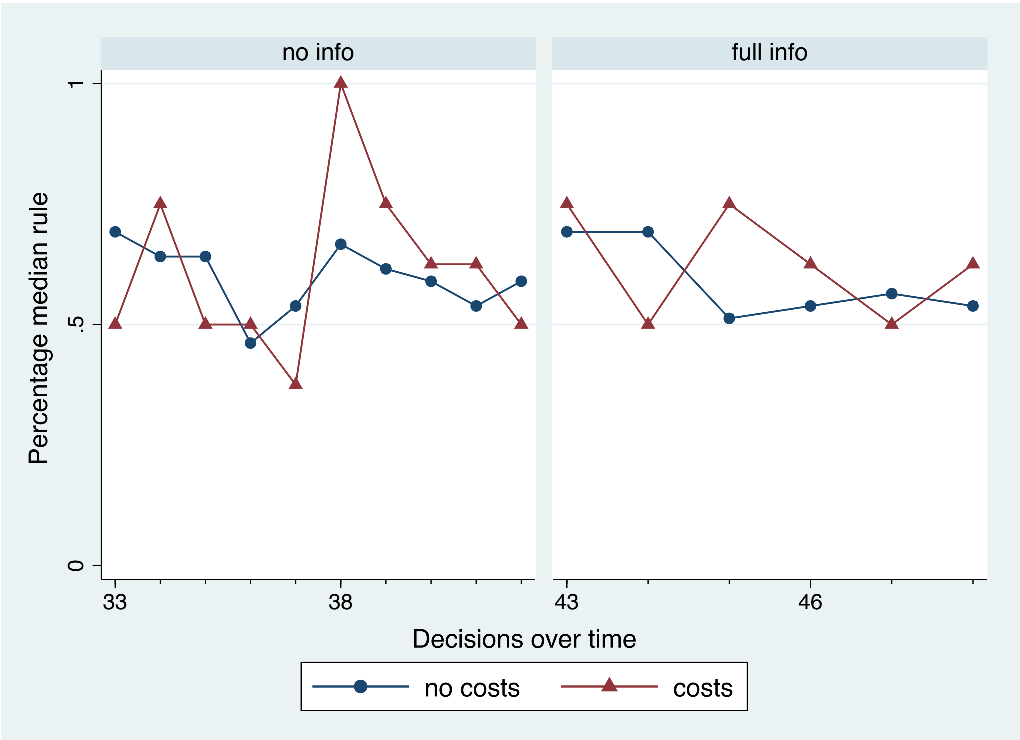


Figure 18.2: Percentage of Median decisions, in which the median of the announced values equals the median of the peaks

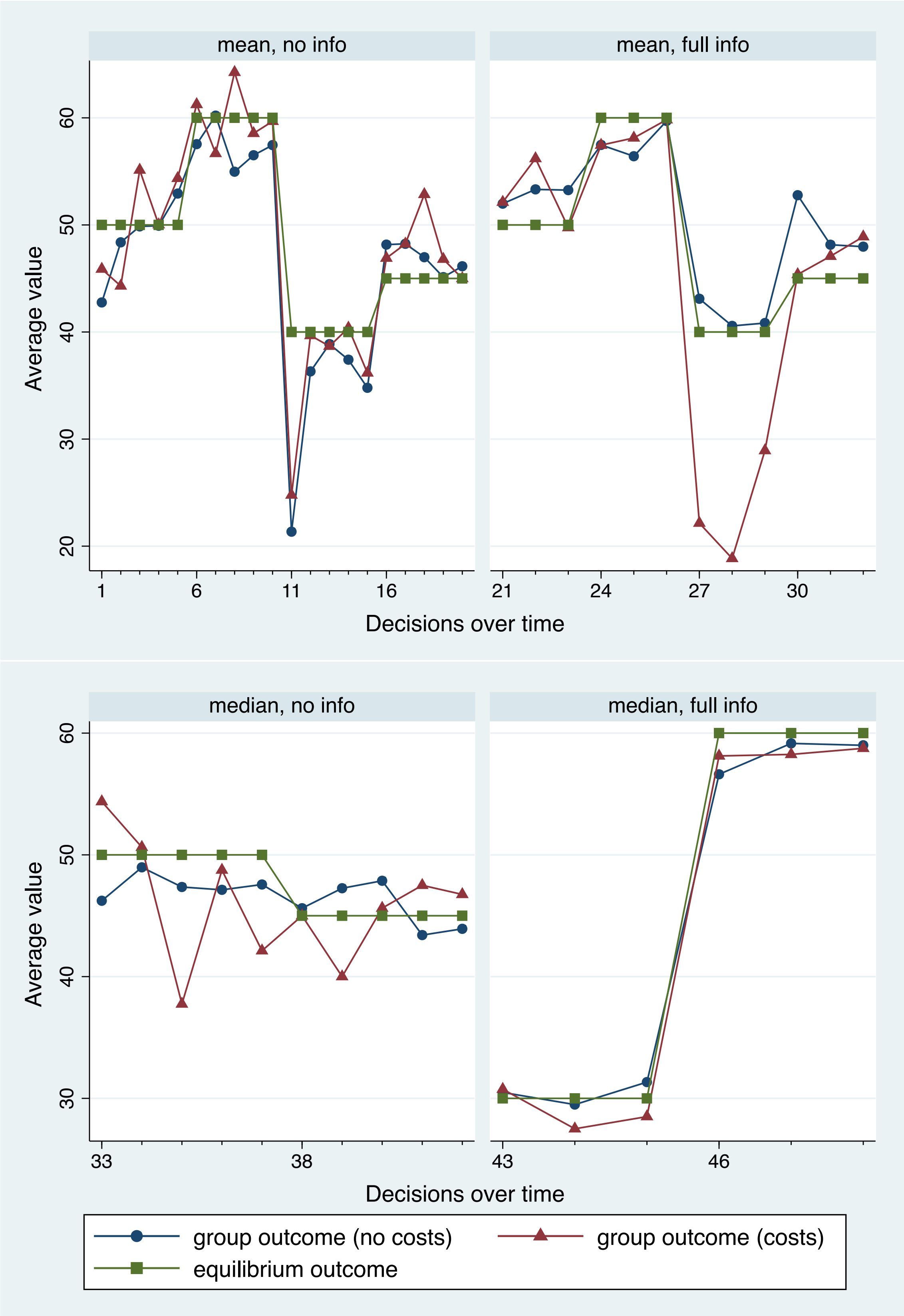


Figure 18.3: Actual outcome and Nash value compared

without costs. With manipulation costs, the group outcome differs a lot from the optimal value in decision 27. We found that the participants with the highest peak (50), instead of telling the high value 100, announced a small value. Up to now, we do not have any explanation for this phenomenon, but recall that the sample size is only four.

We conclude, that under no information (decisions 1-20), the mean value in the first periods differ more from the equilibrium outcome than in following periods, i.e. there is a learning effect. This effect remains present in a weakened form for decisions under full information (decisions 21-32).

For the median rule there are no such outliers and overall the group outcome is closer to the equilibrium when there is full information available (decisions 43-48) than when it is not (decisions 33-42).

18.1.3 Achieving the group Nash play

In Section 18.1.1, we considered whether an outcome coincides with the predicted outcome if every individual acts according to the Nash play. Now, we go a step further and look whether within a group all five individuals play the Nash strategy. These percentages are shown in Table 18.2.

		no costs	*costs*
mean	*no info*	2.8	0
	full info	3.4	0
median	*no info*	2.3 (58.7)	8.8
	full info	3.9 (52.6)	12.5

Table 18.2: Percentage of Nash play by all five group members

Percentages listed here refer to the unique strong straight Nash Equilibrium (see Definition 2). Numbers in brackets consider all straight Nash equilibria, i.e. those where only the individual on rank 3 has to tell the truth and the remaining stay on "their side" of the median peak. Percentages of all straight Nash equilibria are significantly higher compared to only strong straight ones. In the mean treatment only a small percentage of 3% achieves the unique Nash equilibrium under the mean rule when there are no manipulation costs. With manipulation costs not even one group achieved a Nash equilibrium under the mean rule. This supports the conjecture that by introducing costs the calculation of the Nash play becomes a complex task which is not manageable within the short time of the experiment. Under the median rule with manipulation costs, the percentage of pure truth-telling groups rises to 8.8% (no info) and 12.5% (full info).

18.1.4 Efficiency

We saw that on average the social outcomes are close to the theoretically predicted values. In this section we focus on outliers generating an inefficient outcome. Within our context, the definition of Pareto (in-)efficiency reads as follows. A social outcome is inefficient, if it is not within the convex hull of the peaks, i.e. $\bar{x} < x^*_{[1]}$ or $\bar{x} > x^*_{[5]}$.

In total, 117 of 2192 decisions were inefficient, which are 5.3%. Table 18.3 shows the detailed distribution over the experiments and different voting rules. There are more inefficient outcomes under the mean rule than under the median rule. This is not surprising, as for an inefficient outcome under the median rule it is necessary that three or more individuals act irrationally in the same direction at the same moment. Furthermore, for the mean rule the optimal individual action is often outside the convex hull of the peaks but the equilibrium is not, i.e. $x_i^N \notin [x^*_{[1]}, x^*_{[5]}] \ni x^N$. Interestingly, inefficient outcomes are not uniformly distributed: 14 observations were made in decision number 27. Here, we have peaks $(10, 20, 30, 40, 50)$ and the outcome was too high if there were no costs (range from 52 to 67) and too low with manipulation costs (range from 6 to 9), compare also Figure 18.3.

		mean		*median*		all rules
		no info	*full info*	*no info*	*full info*	
no costs	exp. 1	17	24	5	2	48
	exp. 2	26	29	0	0	55
costs	exp. 3	4	5	0	0	9
	exp. 4	3	2	0	0	5
	sum	**50**	**60**	**5**	**2**	**117**

Table 18.3: Inefficient outcomes (total numbers)

Under the median rule with full information 0.7% (and 1.1% with no information) of the outcomes are inefficient. We want to highlight that not even one of them is an equilibrium. In the implementation literature, inefficient equilibria are discussed as a major problem of the median rule. But our results suggest that inefficiency is in fact a negligible phenomenon empirically. Furthermore, all seven inefficient median outcomes were observed in exp.1, where participants already had experience with the mean rule. Here, the sequencing may be an issue.

18.2 Individual aggregated values: an overview

So far, we looked at the aggregated outcome. This is certainly a key figure for many election results. However, in this experiment we were also interested in determining how the result is achieved. Therefore, we look at different strategies under the mean and the median rule and compare truth-telling versus the Nash play.

	truth-telling	Nash play	N°obs
exp 1	0.23	0.62	4560
mean NI	0.12	0.46	1900
mean FI	0.06	0.56	1140
med NI	0.46	0.87	950
med FI	0.56	0.84	570
exp 2	0.24	0.64	4800
med NI	0.47	0.91	1000
med FI	0.54	0.91	600
mean NI	0.13	0.49	2000
mean FI	0.09	0.52	1200
exp 3 (costs)	0.39	0.42	800
med NI	0.62		200
med FI	0.69		120
mean NI	0.21	0.31	300
mean FI	0.24	0.19	180
exp 4 (costs)	0.45	0.47	800
mean NI	0.31	0.32	300
mean FI	0.25	0.31	180
med NI	0.66		200
med FI	0.78		120
Total	0.26	0.60	10960

NI=no information, FI= full information

Table 18.4: Overview over truth-telling and Nash play in all experiments

Some important results can be derived from Table 18.4 (or more compact in Table 18.5). For each experiment and each voting rule, the percentages of truth-telling and Nash play are listed. Each decision is counted independently as one action. The distinction between experiment 1 and 2 is the sequence of the treatments (Mean-Median vs. Median-Mean). In experiments 3 and 4 participants were confronted with manipulation costs. Obviously, the check whether a value coincides with the true peak, is the same for every voting rule, i.e. $(x_i \stackrel{?}{=} x_i^*)$. The "Nash play" refers to the theoretically predicted equilibrium and therefore differs depending on the voting rule and the experiment:

- For the mean rule, there exists a unique equilibrium when there are no manipulation costs, see Section 14.3.1.

	truth-telling	Nash play	N°obs
no costs			
mean NI	0.12	(0.48)	3900
mean FI	0.08	0.54	2340
median NI	0.47	0.89	1950
median FI	0.55	0.88	1170
with costs			
mean NI	0.26	(0.31)	600
mean FI	0.25	0.25	360
median NI	0.64		400
median FI	0.74		240

Table 18.5: Overview over truth-telling and Nash play in all eight treatments

- If there are manipulation costs, the mean rule equilibrium may not exist. These peak distributions are ruled out.[36] Hence, for the mean rule, the announced values are compared to the values given in Table G.1.
- For the median rule without manipulation costs, all choices are taken into account which are a best response if the other players act as in the straight Nash equilibrium, i.e. tell the truth. In particular, we count a value as "Nash play", if the individual is the median player or if he announces a value which is on the same side of the median peak as his own peak.[37] However, truth-telling is a weakly dominant strategy and the corresponding equilibrium is the only SSNE according to Definition 3.
- For the median rule with manipulation costs, the best response coincides with the truth-telling strategy as we analyzed in Section 14.2.1.

Summarizing Table 18.4 suggests the following:

1. Truth-telling is an action which is played more often under the median rule than under the mean rule. This result is robust as it holds as well in each experiment as on the overall consideration. In particular, it is independent of the sequence (Mean-Median) and manipulation costs.
2. Without costs (Exp.1 and 2), the percentage of Nash play is higher under full information than under no information for the mean rule.
3. Without costs, it seems that it does not have an influence whether participants start with the median or mean rule.

[36]See Appendix N for a table where these peak distributions are ruled for the entire analysis.

[37]$(x_{[i]} \overset{?}{\leq} x^*_{[3]})$ for $i = 1, 2$, $(x_{[3]} \overset{?}{=} x^*_{[3]})$ and $(x_{[i]} \overset{?}{\geq} x^*_{[3]})$ for $i = 4, 5$.

4. The percentage of truth-telling under the median rule is higher under full information than under no information.
5. The percentage of truth-telling under the median rule is higher in the presence of manipulation costs.
6. The percentage of truth-telling under the mean rule is higher in the presence of manipulation costs.
7. The percentage of finding the Nash play without costs is similar for mean and median rule.
8. The percentage of finding the Nash play with costs is higher for the median than for the mean rule.
9. The percentage of Nash play under the mean rule is smaller in the presence of manipulation costs.

18.3 The median rule

After having seen a brief overview of the results, we now have a closer look on the two rules. We start with the median rule and compare the announced values with the peaks, i.e. check who tells the truth.

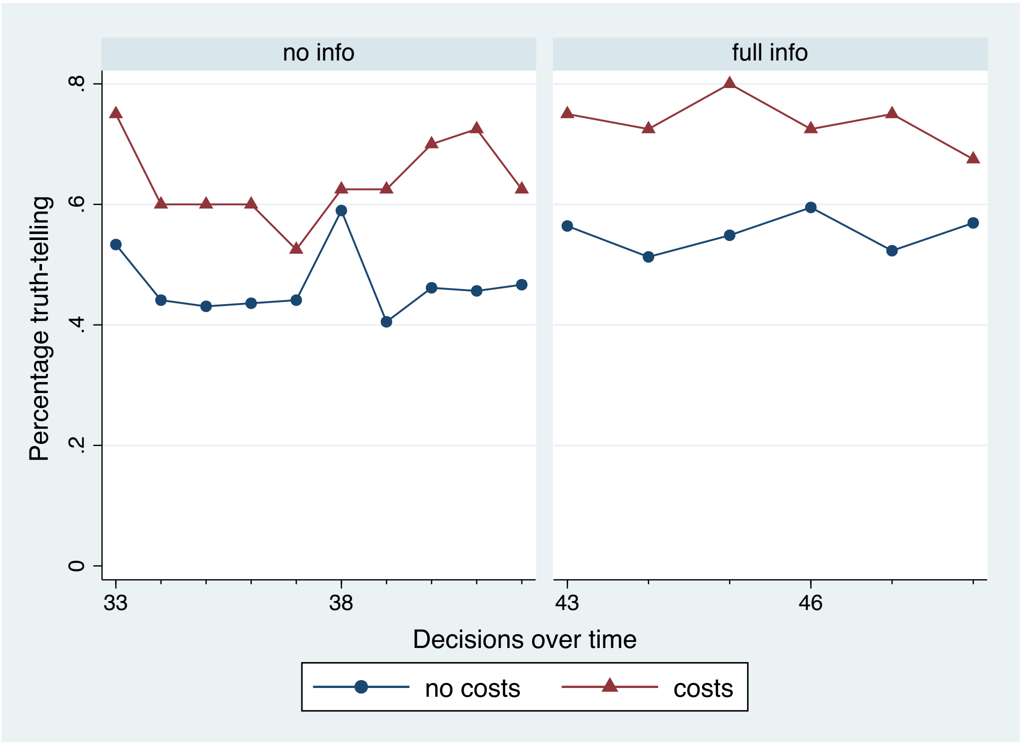


Figure 18.4: Percentage of truth-telling under the median rule

Figure 18.4 shows the percentage of truth-telling under the median rule. The blue *no costs*-line is below the red *costs*-line for each decision, i.e. truth-telling in general occurs more often when there are manipulation costs. Without costs, we observe that truth-telling is higher in the first periods (decisions 33 and 38) than in the following four rounds of each peak distribution when there is no information about the peak distribution available. When full information is available (and three periods are played), the percentage of truth-telling is minimal in the second round.

The percentage of Nash play under the median rule is shown in Figure 18.5. When there are no costs (left subfigures), we can distinguish between two kinds of Nash play: those, which lead to the strong straight Nash equilibrium (see Def. 3) and those which lead to a weak straight Nash equilibrium (see Def. 4). Together, they sum up to the straight Nash play with 89% (no info) and 88% (full info) on average according to Table 18.5. Without information, strong straight and weak straight,

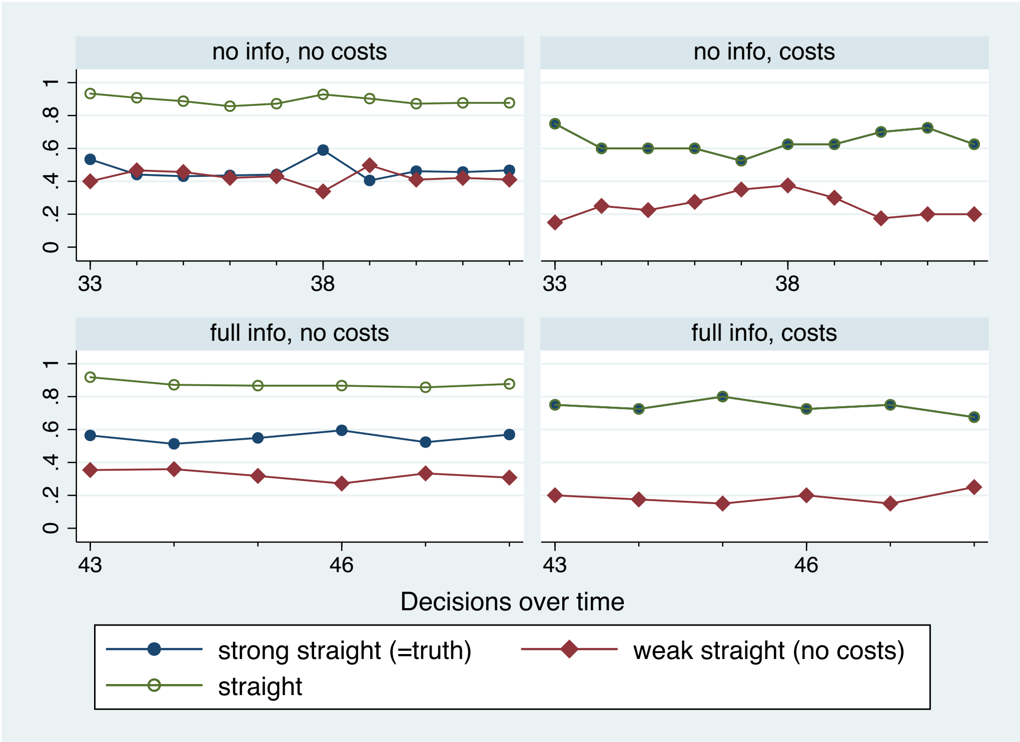


Figure 18.5: Percentage of Nash play under the median rule

i.e. truth-telling and optimal but not truth, are in similar ratio. With information, we observe a shift from weak straight to strong straight.

On the right subfigures, the strong straight Nash play equals the straight Nash play which is truth-telling. The red diamond line shows which plays would have been a weak straight Nash play if there had been no costs. It is only meant as a reference as there does not exist any weak straight Nash play under the median rule when manipulation costs occur.

On average, there are 64% (no info) and 74% (full info) of straight Nash play with costs.

18.4 Other strategies under the median rule: Trying to "win"

We observed that some individuals tend to vote for an intermediate value, which suggests that they either have a preference for being the one who exactly announce the aggregated outcome or for finding a consensus. We distinguish between a

1. *Nash-winner*, who announces the value which is the outcome of the strong straight Nash equilibrium, i.e. the peak of individual with rank 3
2. *winner*, who announces the value, which equals the actual group outcome
3. *consensus-winner*, who announces the value which was played in the previous round. This concept makes sense in periods ≥ 2.

Table 18.6 gives an overview of the percentage of individuals who did not tell their true peak and fit in one of the categories mentioned above. They are sorted by rank. Hence, for rank 3 (=median voter) obviously there can be no Nash winner.

		no costs		costs	
		no info	full info	no info	full info
Rank 1	Nash winner	2.59 %	11.81 %	7.50 %	10.00 %
	winner	11.48 %	21.53 %	10.00 %	10.00 %
	consensus winner	5.41 %	10.00 %	5.88 %	0.00 %
Rank 2	Nash winner	6.07 %	7.89 %	0.00 %	16.67 %
	winner	16.19 %	20.18 %	0.00 %	33.33 %
	consensus winner	7.18 %	2.60 %	8.33 %	0.00 %
Rank 3	Nash winner	0.00 %	0.00 %	0.00 %	0.00 %
	winner	43.04 %	26.09 %	47.06 %	50.00 %
	consensus winner	6.45 %	11.76 %	6.67 %	0.00 %
Rank 4	Nash winner	8.59 %	15.69 %	10.53 %	9.09 %
	winner	21.72 %	23.53 %	21.05 %	27.27 %
	consensus winner	8.72 %	7.35 %	6.67 %	11.76 %
Rank 5	Nash winner	6.48 %	9.22 %	0.00 %	5.88 %
	winner	13.77 %	29.79 %	15.79 %	23.53 %
	consensus winner	6.83 %	13.27 %	0.00 %	18.18 %

Table 18.6: Percentage of "winners." Only non truth-telling decisions are considered

18.5 The mean rule

Under the median rule, independently of the other treatment variables, at least 40% of the announced values equal the peak in each decision. When considering the mean rule, the fraction of truth-telling is significantly lower as we see in Figure 18.6. It shows the percentage of truth-telling under the mean rule.

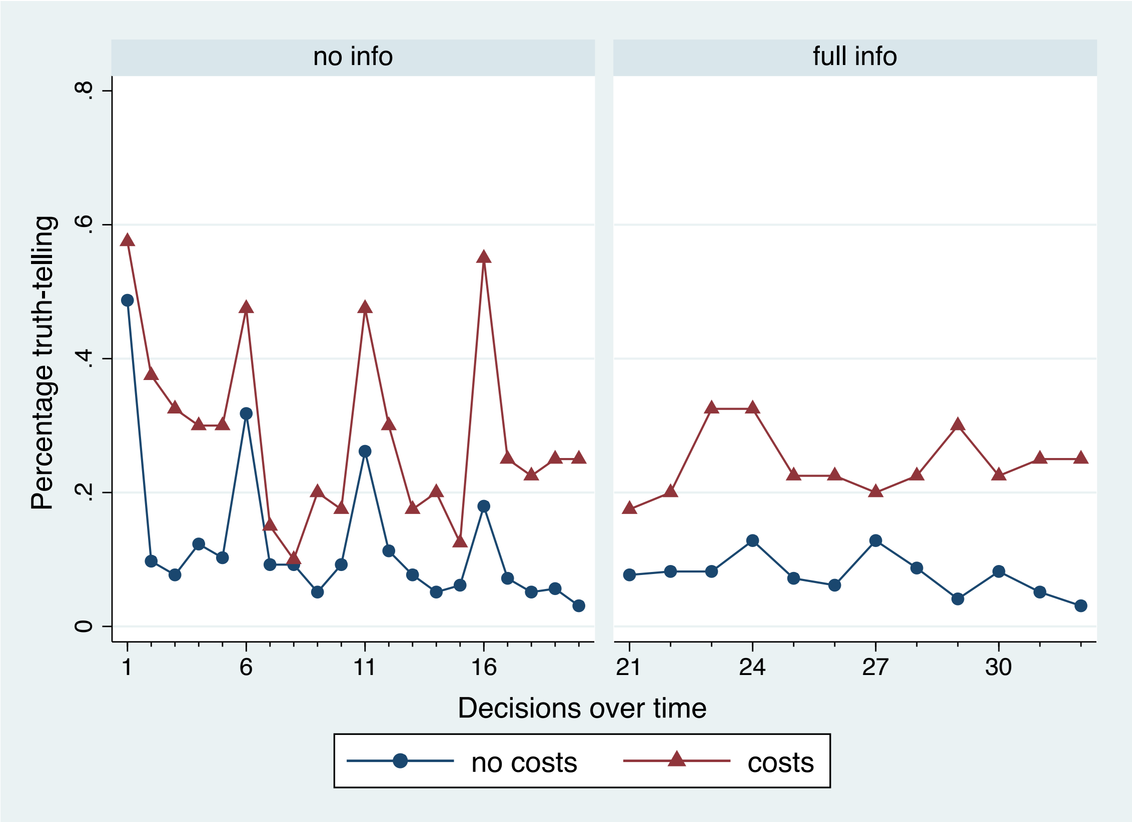


Figure 18.6: Percentage of truth-telling under the mean rule

For each of the 32 decisions the fraction of persons who announce their peak is determined. As in the previous sections, every first period of a peak distribution is labeled with the corresponding decision number. In the *no info* treatment (left subfigure) the functions are spiky whenever a new peak distribution was allocated to the participants. In the *full info* treatment (right subfigure) this observation cannot be made. As the theory predicts, in each decision, the percentage of truth-telling is higher when there are manipulation costs (red line) than when there are no costs (blue line). This result holds for both full and no info. On average, when there are no costs 12% (8%) of the no info (full info) decisions are honest. For manipulation costs, even 26% under no info and 25% under full info are telling the truth.

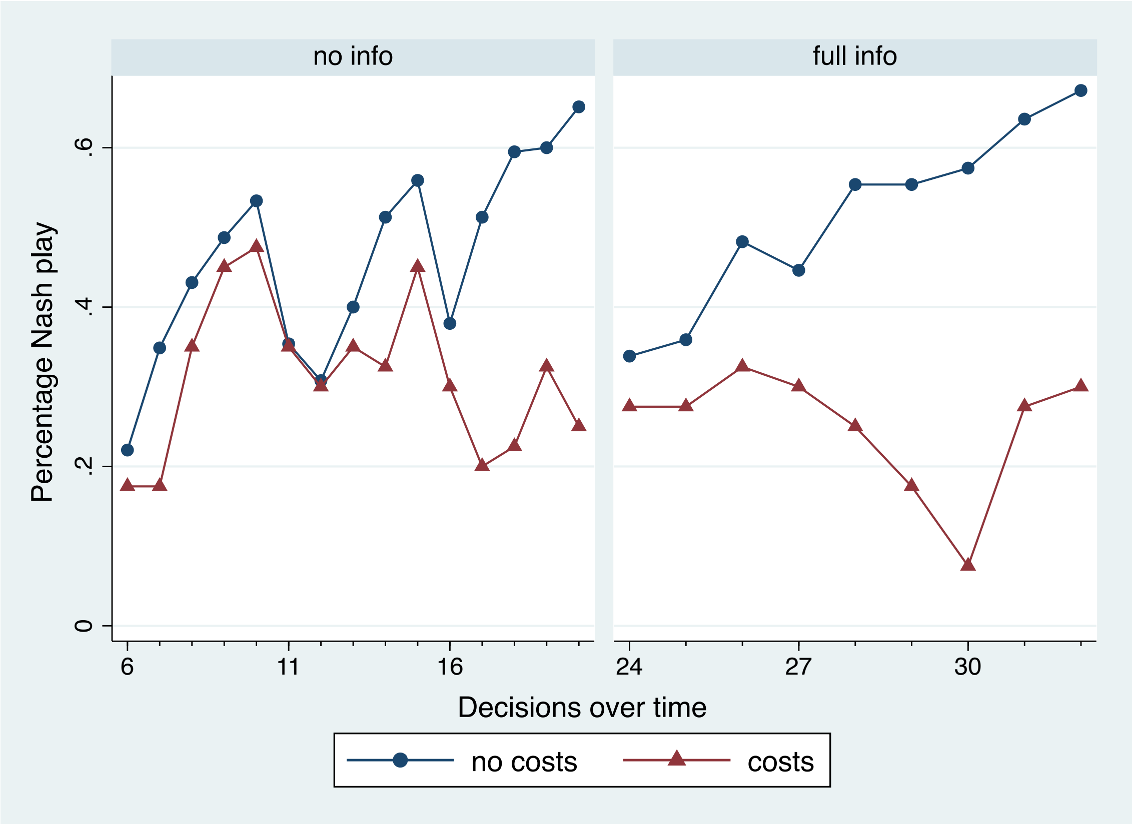


Figure 18.7: Percentage of Nash play under the mean rule

Figure 18.7 shows the percentage of Nash play under the mean rule over time.[38] The left subfigure shows the result under no information. The average values for no costs (costs) are 46% (31%) when no information about peaks is available. Recall that these numbers should only be taken as a reference as it is not possible for the participants to calculate the equilibrium strategy. The three dips of the blue line represent the change of the peak distribution every five periods. This corresponds to the peaks of truth-telling in Figure 18.6.

The right subfigure shows the result under full information. Recall that without costs, a unique Nash equilibrium always exists. Here, the average values for no costs (costs) are 51% (25%) over the appropriate decision numbers. Without costs, there is a tendency towards the Nash play. This learning curve is not affected by the peak distribution, which was changed every third round.

In both info treatments, the *no costs* line lies above the *cost* line.

[38]Due to the non-existence of an equilibrium in the treatment "manipulation costs", only decisions 6 to 20 and 24 to 32 are taken into account.

18.6 Other strategies under the mean rule without manipulation costs

Not only truth-telling and Nash play are strategies that we observed. We focus on three different strategies for mean decisions, namely "best response to true preferences", "best response to previous round results" and "best response to actual choice." These strategies are explained in the following subsections. Due to sample size and the complexity of calculation, we confine ourselves to the experiments without manipulation costs. Table 18.7 shows the percentage of the decisions which fit in the particular strategy definition.

18.6.1 Best response to true preferences

For each decision, it is counted how many decisions would be optimal given that all other players tell the truth. Formally, the calculation reads

$$x_i^* = \frac{1}{5}\left(x_i^{BR_{true}} + \sum_{j \neq i} x_j^*\right)$$

with the restriction that $x_i^{BR_{true}}$ is within the interval $[0, 100]$. This type of response occurs more often than the Nash play. We observe this strategy more often under full information than under no information. See the first column of Table 18.7.

18.6.2 Best response to previous round results

For each decision it is counted how many decisions would be optimal given that the aggregated value of the other players would be the same as in the previous round. So the optimization problem reads

$$x_i^* = \bar{x}^{prev} + \frac{1}{5}(x_i^{BR_{prev}} - x_i^{prev})$$

where $\bar{x}^{prev}$ is the aggregated output of the previous round and x_i^{prev} is the choice of individual i in the previous round. Obviously, the calculation of this value is not possible in all first periods. Between 27% and 51% of the decisions fit in the $x_i^{BR_{prev}}$ concept, but of course many of those values coincide with the extreme values 0 or 100. See the second column of Table 18.7 for exact shares.

Decision	BR true	BR prev. round	BR actual
mean no info	0.43	0.40	0.43
Peak distribution 1	0.46	0.41	0.46
1	0.17	-	0.18
2	0.46	0.33	0.46
3	0.55	0.43	0.57
4	0.56	0.43	0.57
5	0.54	0.45	0.54
Peak distribution 2	0.33	0.34	0.36
6	0.17	-	0.18
7	0.31	0.25	0.31
8	0.37	0.35	0.40
9	0.38	0.36	0.44
10	0.41	0.39	0.48
Peak distribution 3	0.38	0.39	0.38
11	0.29	-	0.19
12	0.32	0.27	0.28
13	0.37	0.35	0.39
14	0.45	0.45	0.49
15	0.45	0.48	0.54
Peak distribution 4	0.54	0.48	0.53
16	0.37	-	0.37
17	0.51	0.44	0.51
18	0.59	0.50	0.56
19	0.61	0.48	0.58
20	0.63	0.49	0.63
mean full info	0.53	0.43	0.49
Peak distribution 5	0.59	0.44	0.54
31	0.58	-	0.52
32	0.62	0.45	0.55
33	0.59	0.44	0.54
Peak distribution 6	0.38	0.32	0.35
34	0.37	-	0.29
35	0.34	0.27	0.34
36	0.42	0.36	0.42
Peak distribution 7	0.51	0.44	0.45
37	0.54	-	0.35
38	0.52	0.41	0.51
39	0.48	0.47	0.51
Peak distribution 8	0.63	0.51	0.60
40	0.62	-	0.54
41	0.63	0.50	0.62
42	0.64	0.51	0.64
Overall	0.46	0.41	0.45

Table 18.7: Percentage of different strategies

18.6.3 Best response to actual choice

For each decision it is counted how many decisions were optimal in the given situation.

$$x_i^* = \frac{1}{5}\left(x_i^{BR_{actual}} + \sum_{j \neq i} x_i\right)$$

This value is hypothetical as it is not calculable by the participants ex ante. As we see in the last column of Table 18.7, the values are higher than the *Best response to previous round* for every decision number.

For a better visualization, the results of the three best response strategies are illustrated in Figure 18.8 for *no info* and *full info* treatments. In many decisions the best response is the same independently of the strategy, e.g. announcing the extreme value 100 could be the best response to true preferences as well as to the results of the previous round. Hence, the plots look very similar.

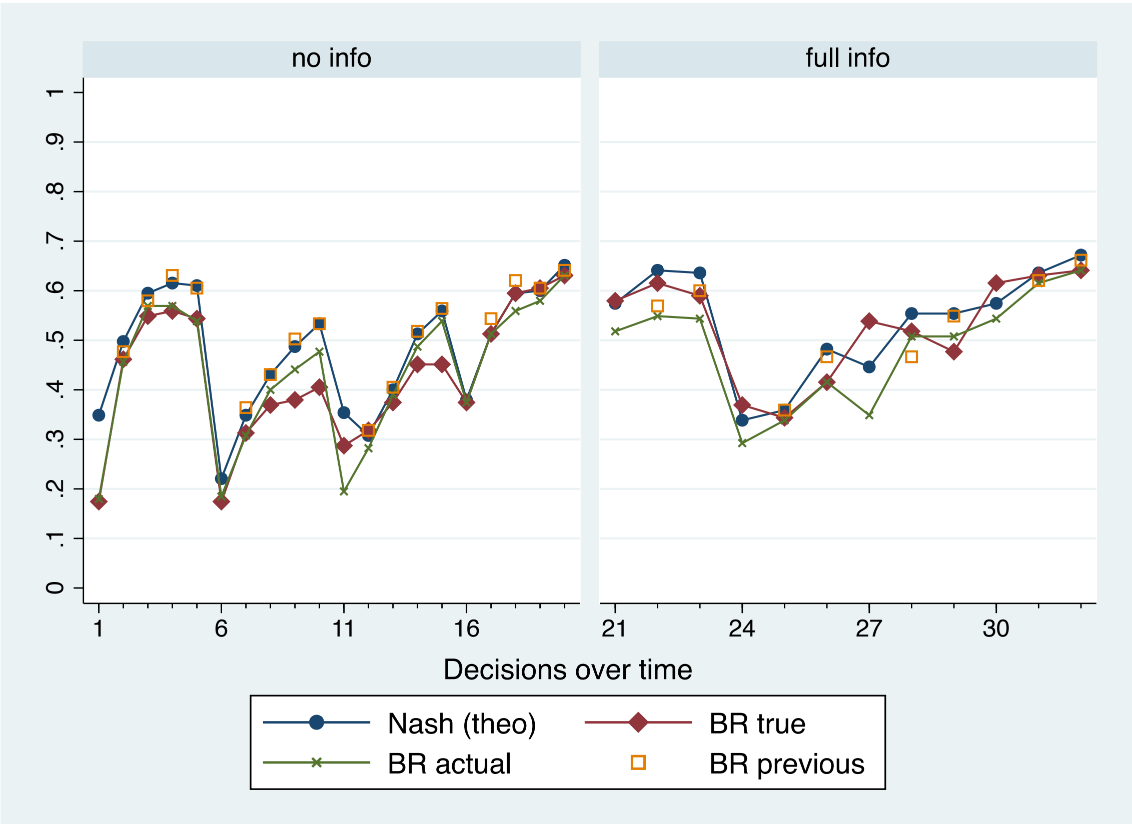


Figure 18.8: Best response functions for the mean rule compared

18.7 Best response to previous round with manipulation costs

In Section 14.3.6 we have analyzed in detail, how the best response function in the presence of manipulation costs looks like. Here, we compare this function with the experimental data for a concrete example, namely for decisions with peak $x_i^* = 20$. Thus, for each observation, we calculate what the mean in the previous round would have been, if the individual with peak 20 had told the truth. Figure 18.9 plots the "mean value in the previous round if telling the truth" against the "actually announced value". Each decision from the experiment is marked by a red circle. The blue line is the plot of the best response function derived in Section 14.3.6.

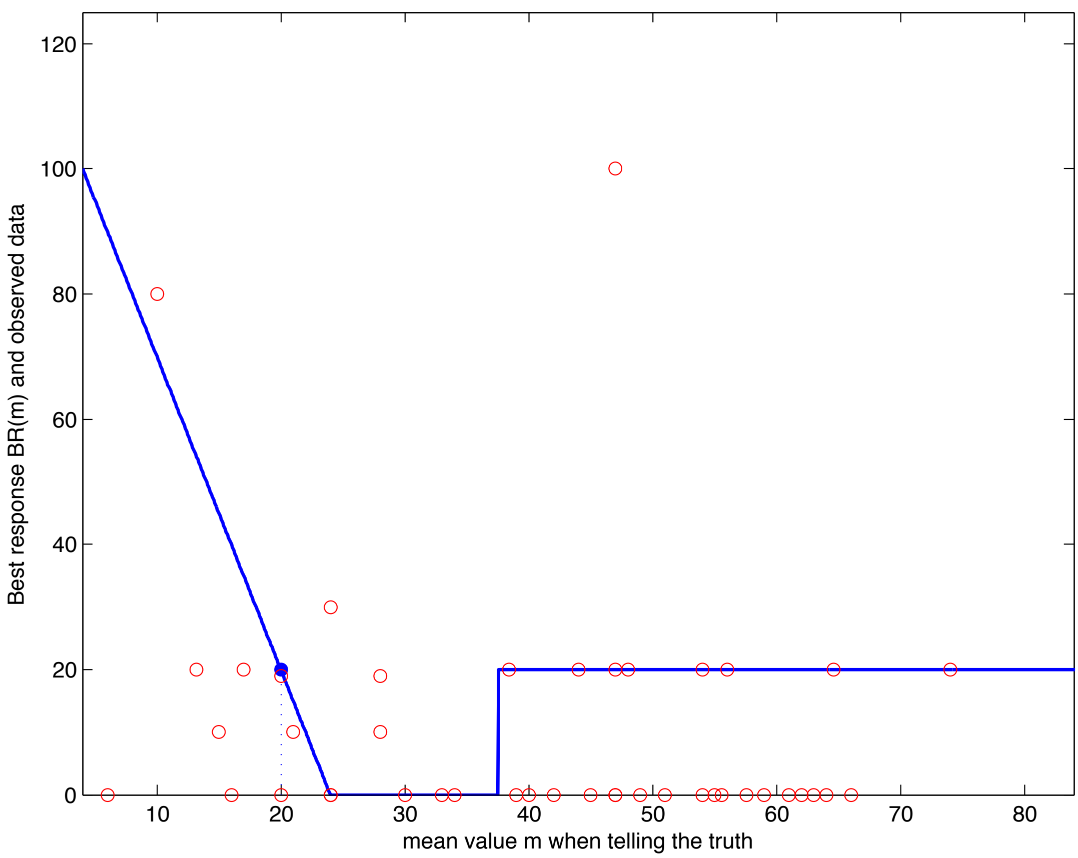


Figure 18.9: Best response to previous round and best response function with manipulation costs for $x_i^* = 20$

We observe that only in one decision a value which is far too high (100) is announced. Most individuals tend to tell either the truth or to manipulate to the extreme value 0. Only three red circles are between 0 and 20, namely at value 10. However, there are many decisions which are according to the best response function, i.e. the red dots lie on the blue function. When the mean value is large, participants announce a value of 0 instead of 20. This implies that they either underestimate manipulation costs or overestimate their influence by manipulating.

18.8 Learning

As we discussed in Section 14.3.2, truth-telling is the same as level-0 learning and *best response to true preferences* can also be interpreted as level-1 learning. Here, we give a short overview over higher levels. The Nash equilibrium is achieved at the latest in level 3.

Table 18.8 gives an overview over the percentage of level-k actions in different mean decisions without manipulation costs. We distinguish between the two info treatments *no info* and *full info*. In each row, the values are aggregated per period, i.e. all first periods (decisions 1, 6, 11 and 16 for *no info*) are aggregated and so on.

The percentage of level-0 in decisions without information is 31% in period 1. Compared to other periods, the level-0 fraction is maximal in this period. In later periods, the percentage of level-1 actions increases, i.e. 40% in period 2 and even 51% in period 5. A reason for this can be that participants update their beliefs. For decisions under full information, the high percentage of level-0 in the first period does no longer hold. We observe 53% level-1 decisions from the first to the third period, i.e. the given information is often used to calculate the level-1 value. We have seen that it depends on information and time, how participants form their beliefs. In the following, we consider the decisions of the participants according to the concrete peak distribution and their player number.

Table 18.9 aggregates the percentages of level-k actions for each distribution per player group. In total, there are data of 195 participants analyzed, i.e. each percentage represents the average value of 39 persons in 5 or 3 decisions depending on the peak distribution. These results can also be compared with the theoretical analysis of Table 14.1, where peak values and level-k values are listed. First, we consider peak distributions with no information (peak distributions 1-4). We observe that for most players, the highest percentage is achieved for the level-k that coincides with the Nash strategy. If the Nash strategy is already the best response to true preferences, i.e. level-1, it is announced more often than in constellations, where it concurs on/with a higher level. As an example, we take a closer look on peak distribution 2: For players 1, 4, and 5, level-1 actions are played with frequency 46%, 71% and 39%, respectively. For those three players, the level-1 action is the announcement of an extreme value (100, 0, 100). In contrast, the Nash strategy for player 3 is reached in level 2 and for player 2 in level 3. The corresponding frequencies are 19% and 27% and therefore less than for players 1, 4 and 5. If the peak is already closer to an

Period	L0?	L1?	L2?	L3?
Mean No Info	**0.12**	**0.43**	**0.46**	**0.48**
Period 1	0.31	0.25	0.31	0.33
Period 2	0.09	0.40	0.40	0.42
Period 3	0.07	0.47	0.50	0.51
Period 4	0.07	0.50	0.54	0.55
Period 5	0.07	0.51	0.57	0.59
Mean Full Info	**0.08**	**0.53**	**0.54**	**0.54**
Period 1	0.10	0.53	0.49	0.48
Period 2	0.07	0.53	0.54	0.55
Period 3	0.05	0.53	0.59	0.59
Overall	**0.11**	**0.46**	**0.49**	**0.50**

Table 18.8: Percentage of level-k actions in mean decisions under no/full information in different periods without manipulation costs ($n = 195$)

extreme value, it is more probable that the player actually announces this extreme value. In peak distribution 4, we observe the smallest value for level-3 for player 4. This is the player whose best response is hard to calculate, as the level-1 response is higher and the level-2 response is smaller than the level-0 value.

Under full information, (peak distributions 5-8), participants take the information about the others' peaks into account. For the level-k analysis, the most interesting constellations are those, were the Nash strategy is obtained for $k \geq 2$. Player 2 in peak distribution 6, for instance, seem to skip level 1 and directly announce the extreme value with a frequency of 25%. For player 6, 15% announce level 0, 10% level 1, 6% level 2 and 3% level 3. On the one hand, this can be interpreted as a process of learning and the existence of different level-k types among the participants. On the other hand, aggregating these frequencies leads to a total of 34% for all level-k strategies, which means that around two third of participants played a different strategy.

Peak Distribution	Player	L0?	L1?	L2?	L3?
1	1	0.39	0.02	0.39	0.39
	2	0.18	0.38	0.38	0.38
	3	0.11	0.62	0.62	0.62
	4	0.10	0.61	0.61	0.61
	5	0.10	0.67	0.67	0.67
2	1	0.12	0.46	0.46	0.46
	2	0.14	0.05	0.02	0.27
	3	0.25	0.03	0.19	0.19
	4	0.04	0.71	0.71	0.71
	5	0.10	0.39	0.39	0.39
3	1	0.07	0.72	0.72	0.72
	2	0.09	0.57	0.57	0.57
	3	0.13	0.13	0.33	0.33
	4	0.11	0.03	0.04	0.08
	5	0.17	0.43	0.43	0.43
4	1	0.08	0.44	0.44	0.44
	2	0.05	0.70	0.70	0.70
	3	0.04	0.79	0.79	0.79
	4	0.04	0.76	0.76	0.76
	5	0.18	0.02	0.04	0.04
5	1	0.03	0.79	0.79	0.79
	2	0.02	0.74	0.74	0.74
	3	0.05	0.54	0.54	0.54
	4	0.29	0.18	0.29	0.29
	5	0.01	0.74	0.74	0.74
6	1	0.03	0.64	0.64	0.64
	2	0.13	0.05	0.21	0.21
	3	0.04	0.72	0.72	0.72
	4	0.09	0.37	0.37	0.37
	5	0.15	0.10	0.06	0.03
7	1	0.09	0.12	0.09	0.17
	2	0.02	0.77	0.77	0.77
	3	0.04	0.72	0.72	0.72
	4	0.03	0.71	0.71	0.71
	5	0.24	0.24	0.22	0.22
8	1	0.04	0.75	0.75	0.75
	2	0.03	0.79	0.79	0.79
	3	0.16	0.09	0.08	0.08
	4	0.03	0.68	0.68	0.68
	5	0.01	0.85	0.85	0.85
	Overall	**0.11**	**0.46**	**0.49**	**0.50**

Table 18.9: Percentage of level-k actions in mean treatments without manipulation costs ($n = 195$)

18.9 (Absolute) Deviation: A comparison of the two rules

So far, we have analyzed the two rules separately and compared whether the announced values were "precision landings" on the strategies truth-telling or Nash play. In this section, we compare the deviation from the peak and the *absolute* deviation from the peak to give a better understanding of the tendency towards a particular strategy. Figure 18.10 shows the average absolute deviation over time. Again, changes of the peak distribution are highlighted by the appropriate label on the horizontal axis. The figure facilitates the comparison of the two rules, both info treatments, and the influence of costs. However, it does not show which individuals deviate and whether the deviation under the mean rule is towards the Nash play.

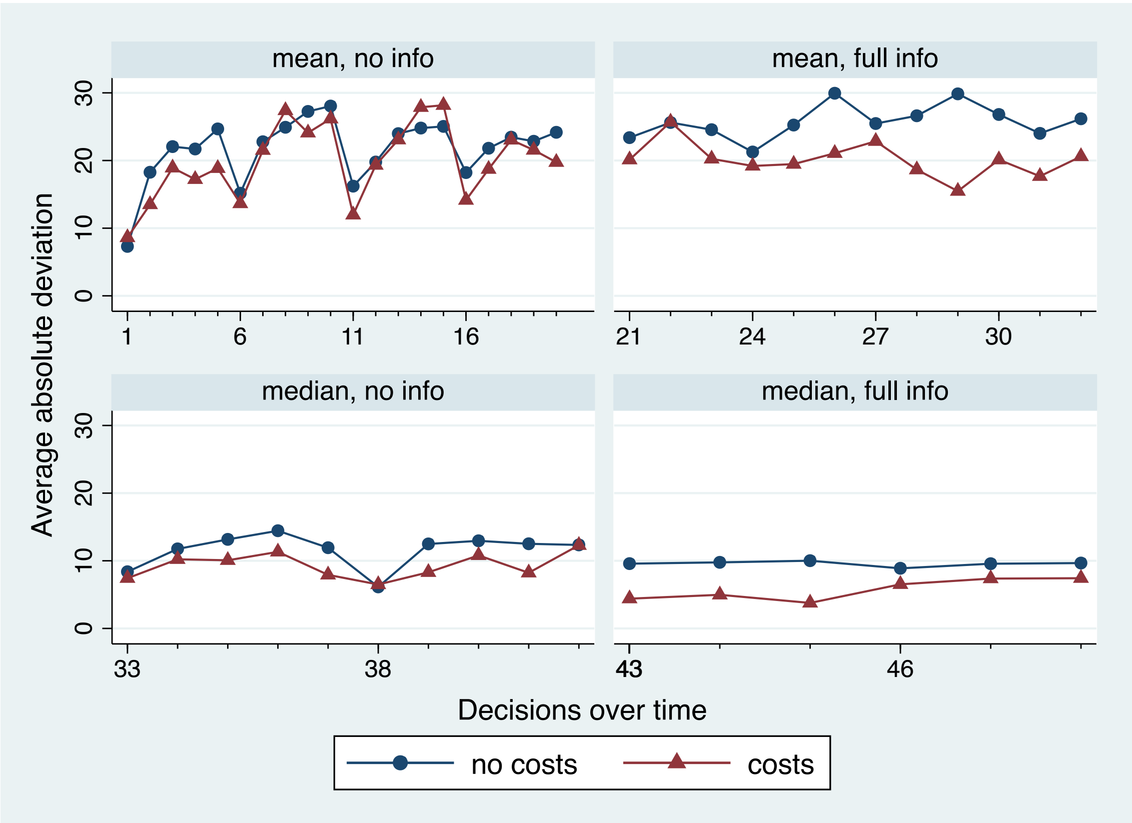


Figure 18.10: Average absolute deviation from the peak over time for decisions with and without costs in different rule and info treatments

Some insight is given in Figure 18.11 where the average deviation (and not longer the absolute deviation) for each rank is illustrated for both cost treatments. The caption of each subfigure consists of the rank (where rank 1 is the player with the smallest peak and rank 5 is the player with the highest peak in this constellation), the rule (mean or median) and the cost (no costs or costs). Within each subfigure, the left bars show the no info decisions and the right bars show the full info decisions.

The first column shows the results for the mean rule when there are no manipulation costs. Especially for the players with rank 2,4 and 5, we see how they manipulate towards the Nash equilibrium. For instance, the average announced values are higher than the peak in each decision for the players with rank 5. There is one exception for rank 4: this is decision number 11, where the new peak 40 was assigned and players deviated to a smaller value instead of a higher value.[39] In the full info treatment this outlier does not occur. For rank 3, it depends on the peak distribution, whether deviation has a positive or a negative sign (see also Table G.1 in the Appendix).

Concerning the median rule, we observe that the average deviation is smaller than under the mean rule. For instance, the player with rank 3 has an average deviation close to zero (third column). Recall that in the cost treatments per bar only eight observations were made and therefore variations are more noticeable.

We sum up that absolute deviation from the peak is higher under the mean rule than under the median rule according to Figure 18.10 and varies more when no information is available. Furthermore, we have seen that the direction and also the strength of the deviation depends on the rank and the available information.

[39] It would be very interesting to know, whether individuals deviate towards the Nash play in this situation when they get to know their rank as an additional information.

Figure 18.11: Average deviation from the peak over time for decisions with and without costs in different rule and info treatments for each rank

18.10 Individually announced values

The previous analysis was based on the aggregation of data over different individuals. Even if data were grouped by rank, an individual belonged to different groups depending on the particular peak distribution. In this section, we discuss whether the tendency for truth-telling is an individual property or whether it is equally distributed among the participants.

18.10.1 Individual truth-telling

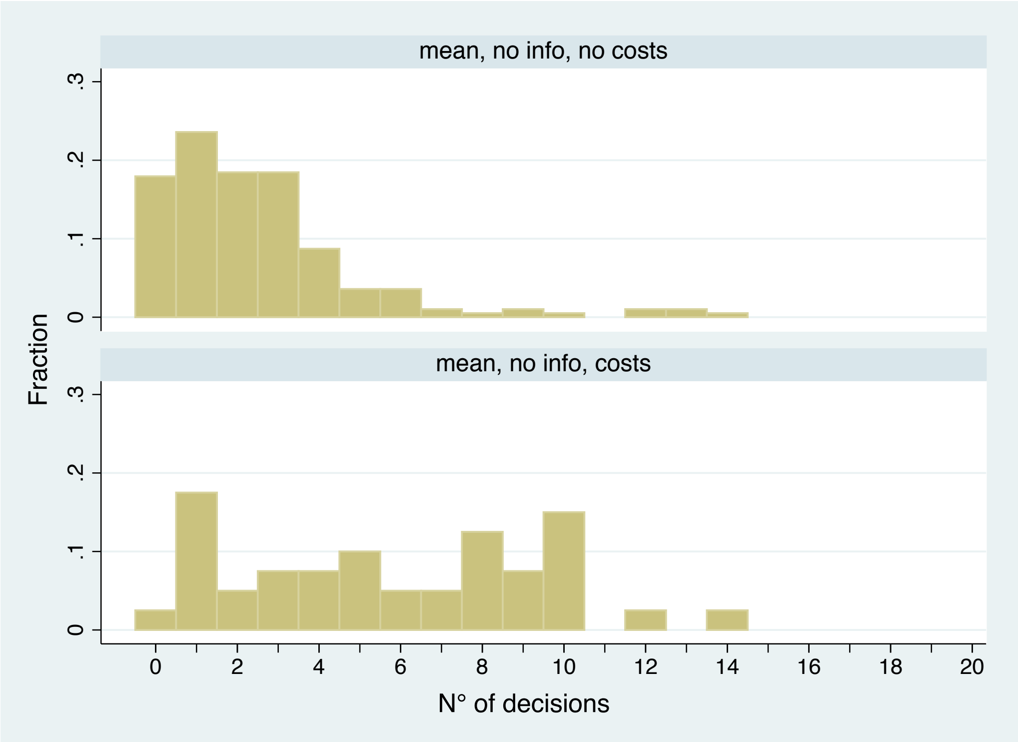
First, we look at the frequency of truth-telling. Figure 18.12 shows the fractions under the mean rule. On the horizontal axis, the total number of decisions is plotted (20 for *no info* and 12 for *full info*). The vertical axis shows the fraction of individuals who told the truth in the corresponding number of decisions. Without costs, 17.95% never announced the peak under no information (Figure 18.12(a)). When manipulation costs occur, the percentage for never telling the truth is only 2.5%.

When full information is available (Figure 18.12(b)), the tendency towards manipulating when there are no costs, is stronger, i.e. 47.69% never tell the truth. With manipulation costs, still 15.00% never tell the truth.

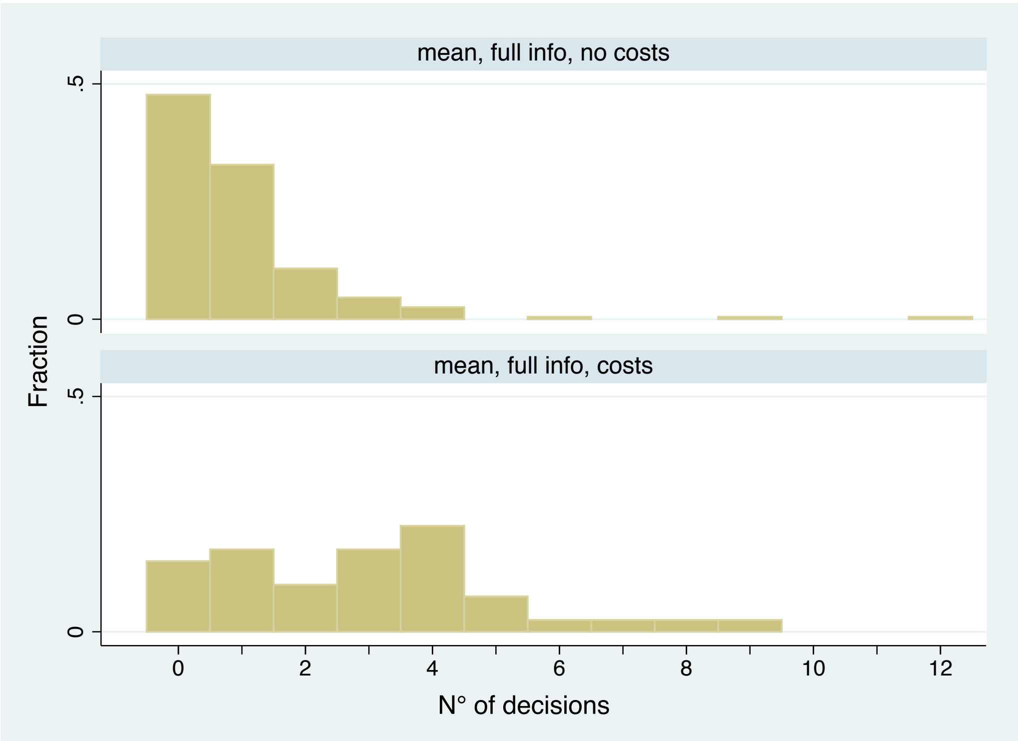
Figure 18.13 shows the fraction of truth-telling under the median rule. Without information (Figure 18.13(a)) and without costs, 9.74% announce the true peak in nine or ten out of ten times and 4.62% never announce the true peak. When there are manipulation costs, the fraction of those manipulating at most once arises to 37.50%. On the other hand, the percentage of never telling the truth is only 2.5%.

Under full information (Figure 18.13(b)), 11.28% never tell the truth when there are no costs and 7.5% when there are manipulation costs. Of the participants, 31.28% manipulated at most once occurs in the no costs treatment, and 60% in the costs treatment.

To sum up, we have seen that under the mean rule without costs, there are some players who never announce their true peak. With costs, the participants announce the truth in at most half of the mean rule decisions. Under the median rule, truth-telling was higher in both cost treatments. But even when there was no incentive to manipulate, we observed individuals who never announce their true peak.

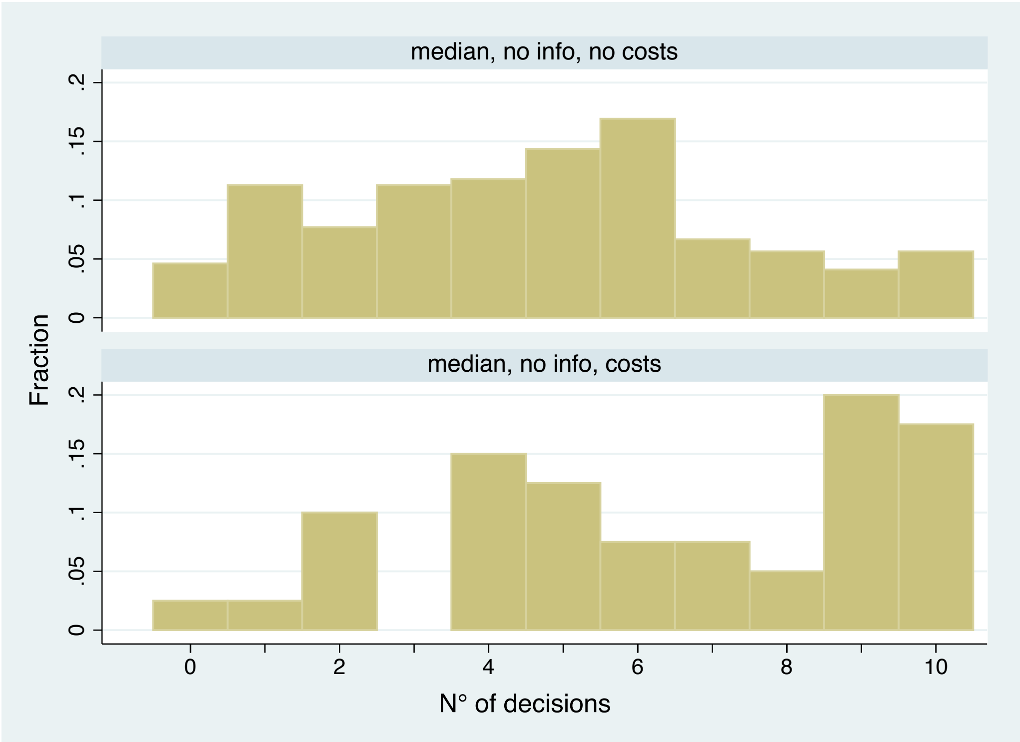


(a) Frequency of individual truth-telling under the mean rule with no information

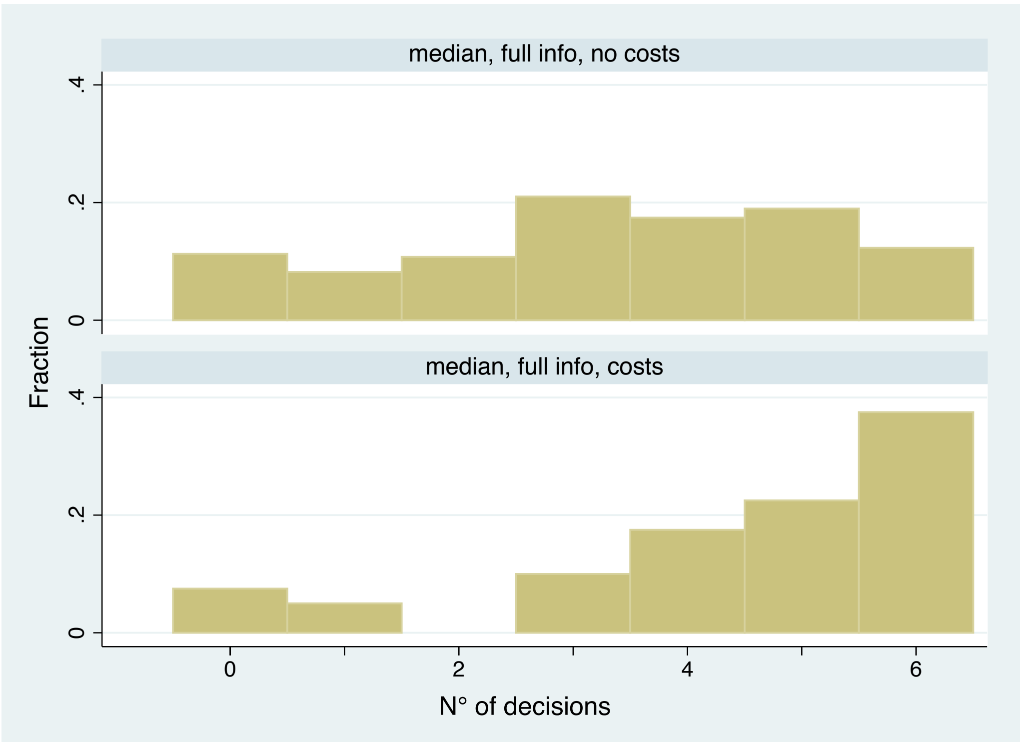


(b) Frequency of individual truth-telling under the mean rule with full information

Figure 18.12: Histograms showing individual truth-telling frequencies under the mean rule



(a) Frequency of individual truth-telling under the median rule with no information



(b) Frequency of individual truth-telling under the median rule with full information

Figure 18.13: Histograms showing individual truth-telling frequencies under the median rule

18.10.2 Individual Nash play and best response

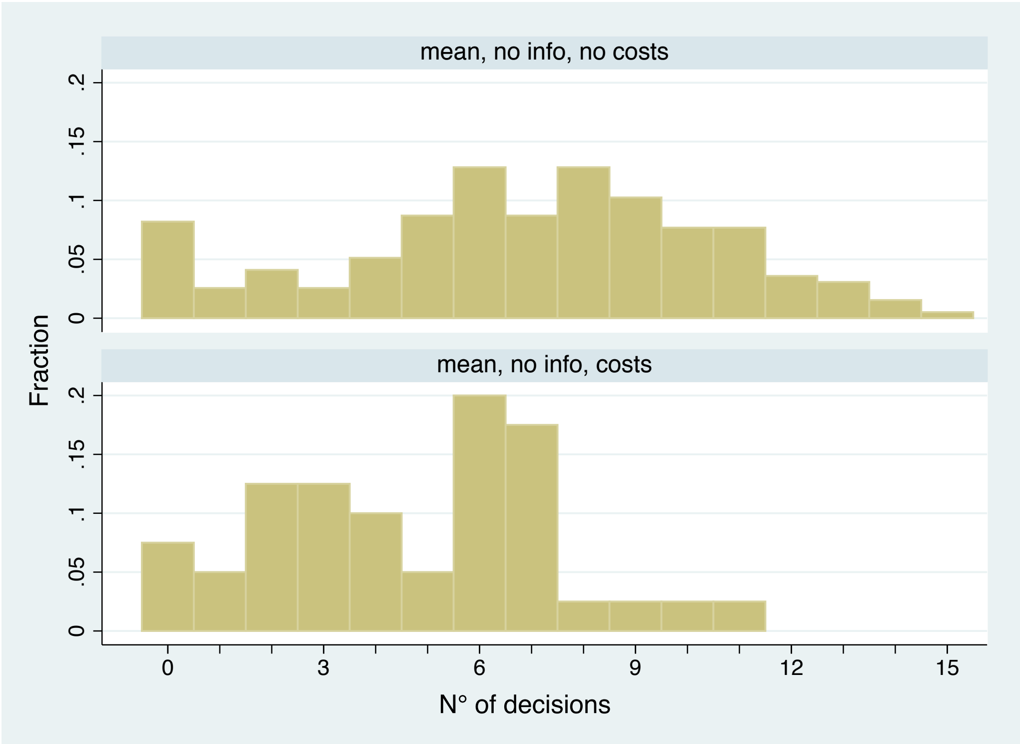
In this subsection we analyze how often the individuals play the Nash strategy. For the mean rule, the results are illustrated in Figure 18.14. In peak distributions 1 and 5, a Nash equilibrium with manipulation costs does not exist (see Appendix H). Therefore, we drop data of this peak distributions for both cost treatments to have better comparability. In the no info treatment 15 of 20 and in the full info treatment 9 of 12 decisions remain, respectively. When there are manipulation costs and no info (Figure 18.14(a)), none of the participants announced the Nash play more often then eleven times and only 10% achieved the Nash play eight or more times. Without costs, this partition is not that strict and each possible number of Nash plays is achieved at least once. The fraction of participants who never achieve the Nash play is similar in both treatments (no costs: 8.21% and costs: 7.5%). With full information (Figure 18.14(b)), the tendency to Nash plays is higher in the no costs treatment. Without costs, 7.18% of the participants announced the Nash play at most once, with manipulation costs, the value rises to 32.5%.

Under the median rule, the number of best responses are shown in Figure 18.15. With costs, the best response equals truth-telling and therefore is identical to Figure 18.13. Without costs, we consider both types of straight equilibria, the strong and the weaker straight ones. Under no info, 78.69% of the participants gave a best response in at least eight of ten decisions. Under full information, even 80.51% of the participants gave a best response in at least five of six decisions.

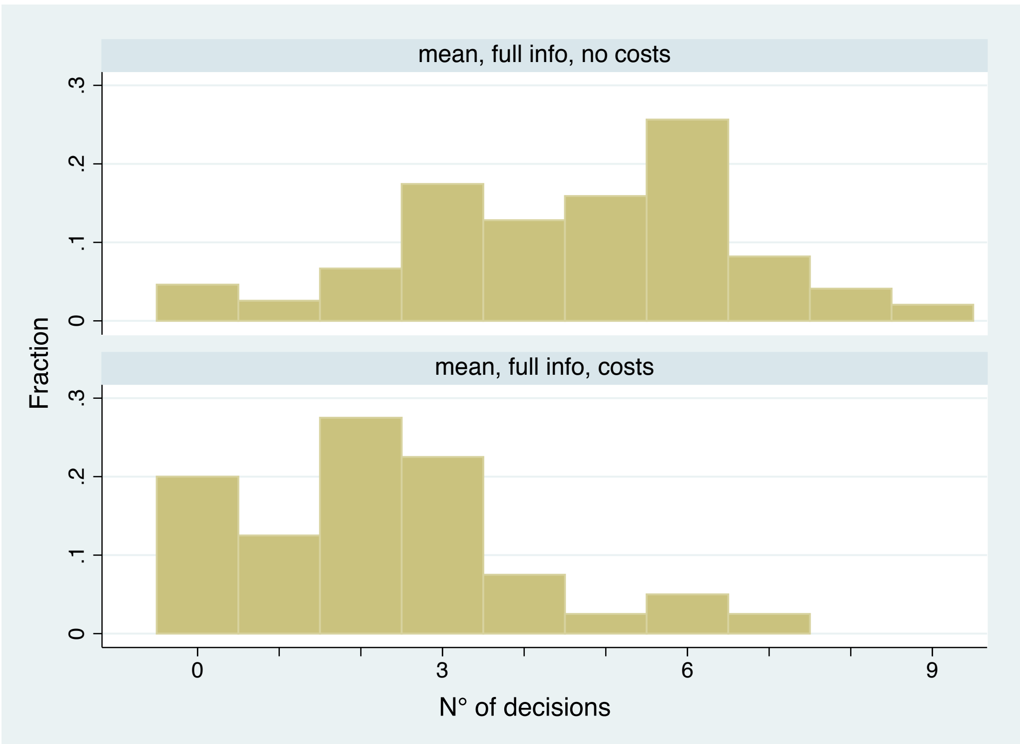
In Figure 18.16 we finally aggregate the individual data over all decisions, i.e. over *rule* and *info*. Interestingly, in both *cost*-treatments, none of the participants announced the true peak in all 48 decisions. Without costs, the true peak is announced between one and 29 times (with one exception at 41). With manipulation costs, individuals tell in five to 38 decisions the truth (see Figure 18.16(a)).

A Nash play (in the sense of a strategy corresponding to a straight Nash equilibrium) was announced in 12 to 38 (of 40) decisions in the treatment without manipulation costs. With manipulation costs, Nash play was observed between 6 and 28 times (see Figure 18.16(b)).

We conclude that the introduction of manipulation costs leads to a shift towards truth-telling and away from Nash play on the individual level.

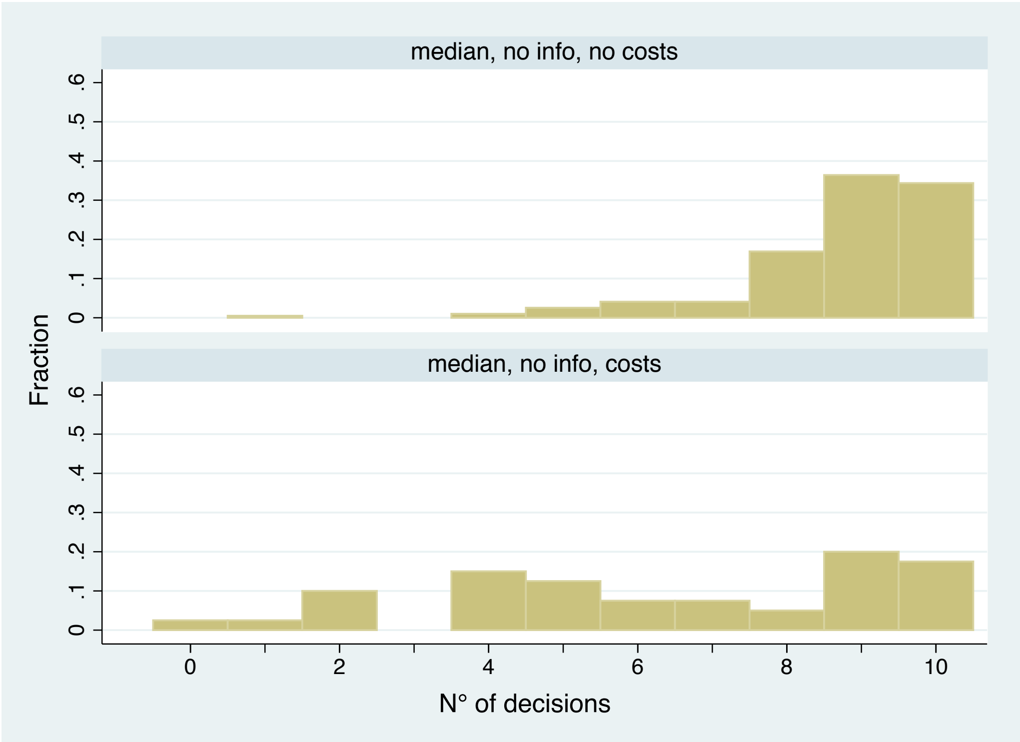


(a) Frequency of individual Nash play under the mean rule with no information for peak distributions 2,3,4

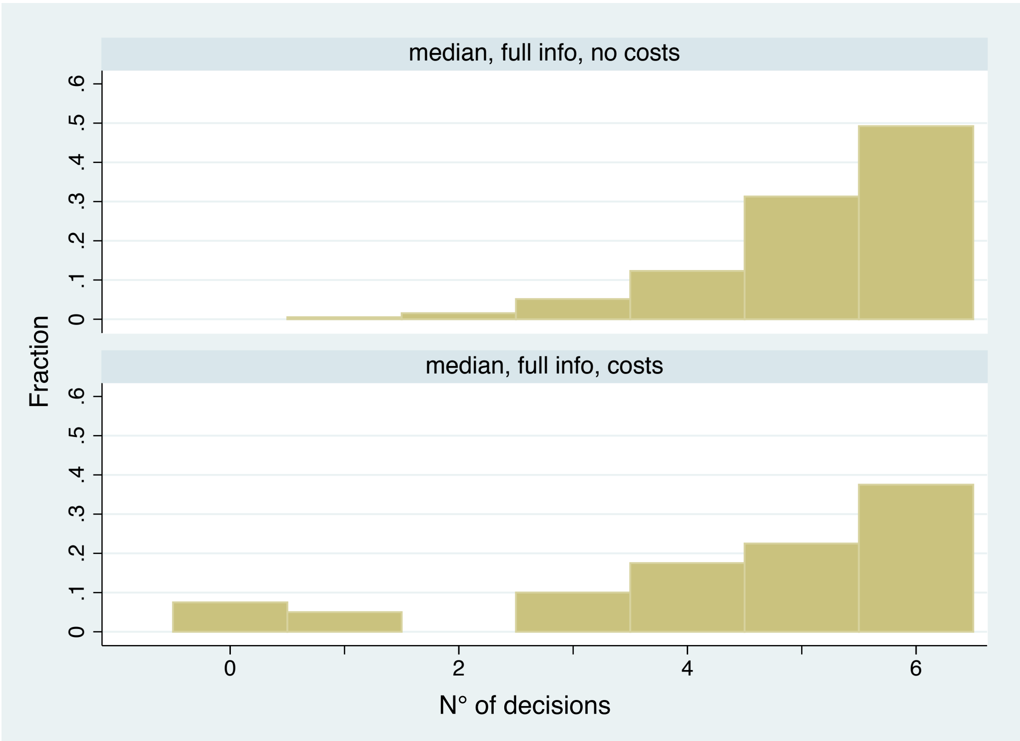


(b) Frequency of individual Nash play under the mean rule with full information for peak distributions 6,7,8

Figure 18.14: Histograms showing individual Nash play under the mean rule

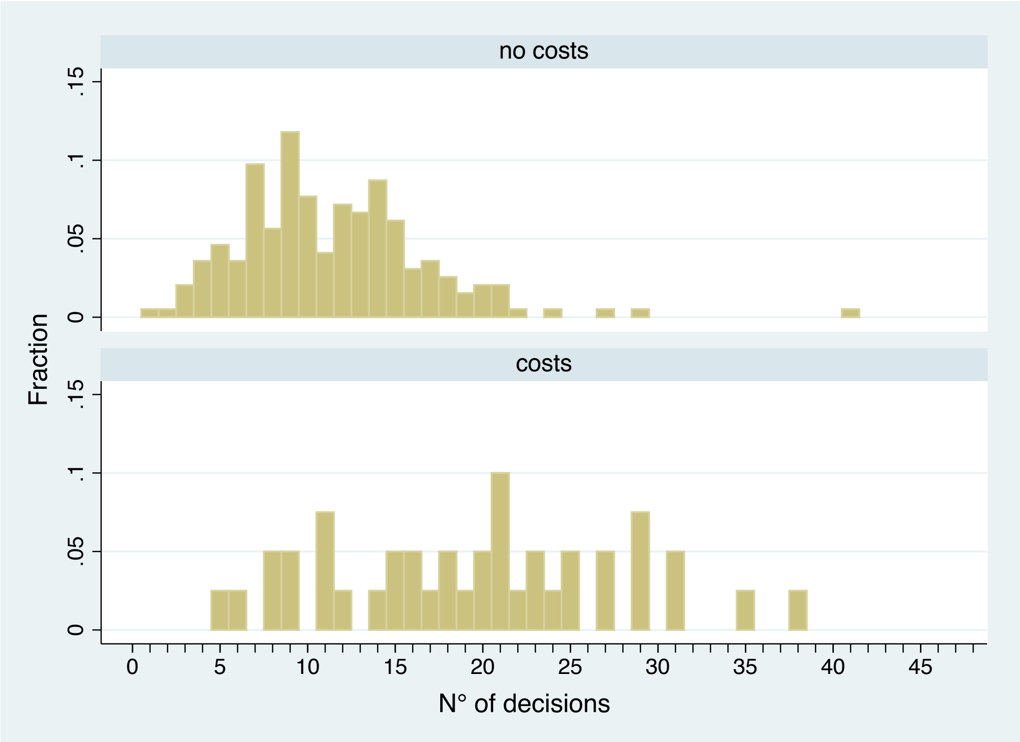


(a) Frequency of individual best response under the median rule with no information

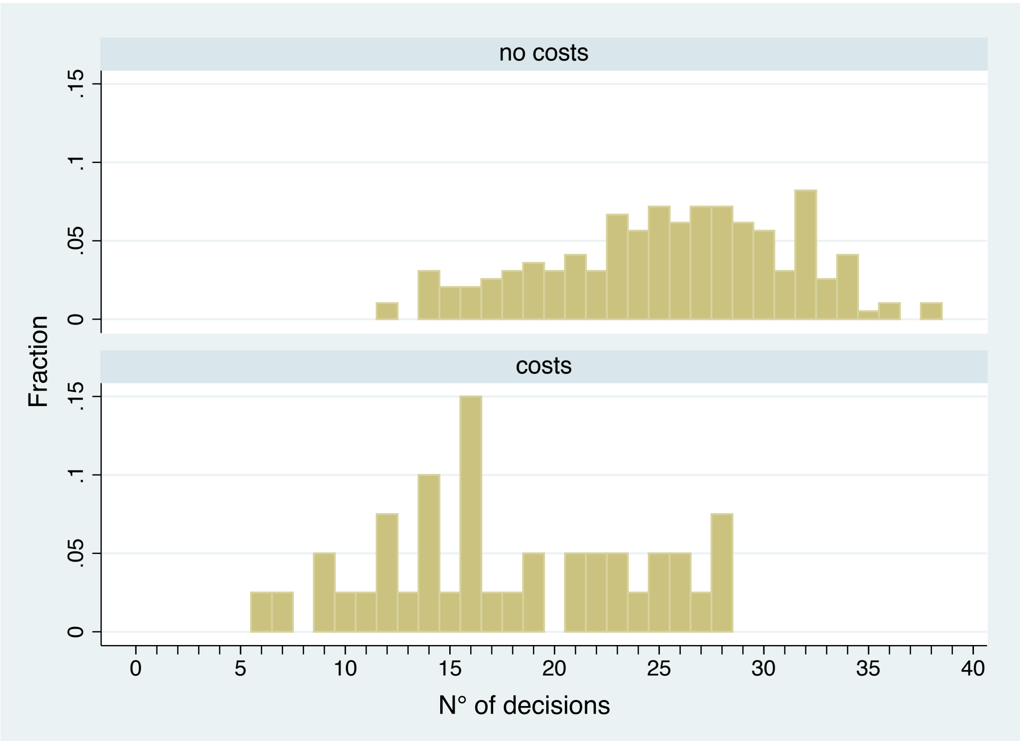


(b) Frequency of individual best response under the median rule with full information

Figure 18.15: Histograms showing individual best response under the median rule



(a) Frequency of individual truth-telling in 48 decisions



(b) Frequency of individual (straight) Nash play in 40 decisions

Figure 18.16: Individual data aggregated over rule and info

19 Experimental results

The previous section gave us an overview over the collected data in a descriptive way. The aim of what follows is a statistical analysis including some hypotheses tests and a regression analysis. In our experiment, interacting groups consisted of five players. Hence, the results of different members of one group are not independent from each other. Therefore, we aggregate the group results and get observations for each group, i.e. 39 independent observations without costs and 8 observations with manipulation costs for each info/rule combination. We have seen that individual deviation from the peak mostly is consistent with the direction of Nash prediction (see Figure 18.11), so we use the average absolute deviation of the group as the dependent variable for our treatment comparisons. The aggregated values are metric variables which suggest to be normally distributed. Hence, parametric tests, such as the t-test can be applied.

In the following, all hypotheses are formulated in a negative way such that we reject them by the appropriate statistical test.

19.1 Influence of "rule"

19.1.1 Influence of "rule" on truth-telling

We claim that truth-telling is higher under the median rule than under the mean rule in all treatments. The H_0 hypothesis reads as follows:

Hypothesis H 1. The average absolute group deviation from the peak is smaller (or equal) under the mean rule than under the median rule in all combinations of cost and info.

To reject these hypotheses for the four treatments which occur when combining "info" (no info, full info) with "costs" (no costs, costs), we run paired t-tests. Table 19.1 and 19.2 show the results: the absolute deviation under the mean rule compared to the absolute deviation under the median rule is signficantly higher ($p<0.0003$) in each treatment.

no costs

Variable	Obs	Mean	Std.	Err.	Std.	Dev.
GroupAbsDev mean	39	21.62	0.46	2.85	20.70	22.55
GroupAbsDev median	39	11.61	0.54	3.35	10.52	12.70
diff	39	10.01	0.74	4.61	8.51	11.51

mean(diff)= mean(GroupAbsDev mean-GroupAbsDev median)
t= 13.55

H_0: mean(diff) = 0, degrees of freedom = 38
Ha: mean(diff) $<$ 0 , Pr(T<t)= 1.0000
Ha: mean(diff) $\neq$ 0 , Pr(|T|>|t)| = 0.0000
Ha: mean(diff) $>$ 0 , Pr(T>t)= 0.0000

costs

Variable	Obs	Mean	Std.	Err.	Std.	Dev.
GroupAbsDev mean	8	19.88	1.21	3.43	17.02	22.75
GroupAbsDev median	8	9.31	1.35	3.81	6.12	12.50
diff	8	10.57	1.74	4.92	6.46	11.69

mean(diff)= mean(GroupAbsDev mean-GroupAbsDev median)
t= 6.08
H_0: mean(diff) = 0, degrees of freedom = 7

Ha: mean(diff) $<$ 0 , Pr(T<t)= 0.9997
Ha: mean(diff) $\neq$ 0 , Pr(|T|>|t)| = 0.0005
Ha: mean(diff) $>$ 0 , Pr(T>t)= 0.0003

Table 19.1: T-Test, no info: absolute deviation from peak is significantly higher under the mean rule than under the median rule.

no costs

Variable	Obs	Mean	Std.	Err.	Std.	Dev.
GroupAbsDev mean	39	25.74	0.53	3.29	24.67	26.80
GroupAbsDev median	39	9.58	0.67	4.19	8.22	10.94
diff	39	16.16	0.84	5.27	14.45	17.87

mean(diff)= mean(GroupAbsDev mean-GroupAbsDev median)
t= 19.15
H_0: mean(diff) = 0, degrees of freedom = 38

Ha: mean(diff) $<$ 0 , Pr(T<t)= 1.0000
Ha: mean(diff) $\neq$ 0 , Pr(|T|>|t)| = 0.0000
Ha: mean(diff) $>$ 0 , Pr(T>t)= 0.0000

costs

Variable	Obs	Mean	Std.	Err.	Std.	Dev.
GroupAbsDev mean	8	20.10	0.94	2.67	17.88	22.32
GroupAbsDev median	8	5.75	1.03	2.90	3.32	8.17
diff	8	14.35	0.96	2.72	12.08	16.63

mean(diff)= mean(GroupAbsDev mean-GroupAbsDev median)
t= 14.94
H_0: mean(diff) = 0, degrees of freedom = 7

Ha: mean(diff) $<$ 0 , Pr(T<t)= 1.0000
Ha: mean(diff) $\neq$ 0 , Pr(|T|>|t)| = 0.0000
Ha: mean(diff) $>$ 0 , Pr(T>t)= 0.0000

Table 19.2: T-Test, full info: absolute deviation from peak is significantly higher under the mean rule than under the median rule.

With absolute group deviation, the number of those individuals who exactly announce the peak cancels out. Therefore, we also run a paired t-test on precise truth-telling.

Hypothesis H 2. The average percentage of truth-telling per group is higher (or equal) under the mean rule than under the median rule in all combinations of cost and info.

These hypotheses can be rejected and the results are highly significant for all treatments ($p<0.0001$).[40]

19.1.2 Influence of "rule" on Nash play

Does the implemented rule have a significant influence on Nash play? We use the theoretical concept of straight Nash equilibria as introduced in Section 14 and run another paired t-test for the percentage of Nash play per group.

Hypothesis H 3. The average percentage of Nash play per group is higher (or equal) under the mean rule than under the median rule in all combinations of cost and info.

This hypothesis can be rejected for all four treatments with $p<0.0001$. The mean of the calculated test variables are similar to the values in Table 18.5 and therefore not listed here.

19.2 Influence of "info"

19.2.1 Influence of information on truth-telling

In this section, we analyze the treatment variable "info". When its value is "no info", the participants do not have any information over the peaks of the other four players. Under "full info", the peaks of the five players are common knowledge.[41] Information about the peaks of the other participants influences the players in two different ways:

1. On the one hand, knowing the peak distribution enables the participants to determine the optimal strategies. Especially under the mean rule, without knowing the own rank, it is not possible to calculate a Nash equilibrium. Hence, full info supports strategic behavior.

[40] Detailed summary statistics are available upon from the author.

[41] In particular, either all players know all peaks or each player only knows his own peak.

2. On the other hand, with information it is possible to draw inferences from the aggregated value about the behavior of the others. Hence, manipulation is more transparent. A fully rational player does not care about this transparency. But there are at least two reasons, why participants care: First, not telling the truth is similar to “lying” which is considered as bad behavior in large parts of our civilization. Second, under the median rule, there is no rational reason to misrepresent the preferences. Many people do not like to be observed when they act irrational.

These arguments sometimes give contradictory recommendations. It is reasonable to assume that the strength of the argument depends on the rule. Under the mean rule, we expect that full information decreases truth-telling, whereas under the median rule, it increases truth-telling.

Hypothesis H 4. The average absolute group deviation is smaller (or equal) under full information than under no information under the mean rule.

We test this hypothesis by a paired t-test.[42] We reject the hypothesis with a p-value of 0.0000 when there are no manipulation costs. With a p-value of 0.4178, we cannot reject the hypothesis when manipulation costs occur. Furthermore, the sample size of eight is too small to make a statement which is statistically significant.

Hypothesis H 5. The average absolute group deviation is larger (or equal) under full information than under no information under the median rule.

We can reject this hypothesis with a p-value of 0.0011[43] (0.0109) when there are no manipulation costs (with manipulation costs). However, the result is only statistically significant for the “no costs” treatment with 39 groups.

We also check whether the precise truth-telling is different for the treatment variable “info”:

Hypothesis H 6. The average percentage of truth-telling per group is higher (or equal) under full information than under no information under the mean rule.

The hypothesis can be rejected with a p-value of 0.0000 for no manipulation costs and the result is not significant for manipulation costs (p=0.0876).

Hypothesis H 7. The average percentage of truth-telling per group is smaller (or equal) under full information than under no information under the median rule.

The hypothesis can be rejected with a p-value of 0.0000 for no manipulation costs and the result is not significant for manipulation costs (p=0.0013).

[42] We run the sdtest for all variables to ensure that a t-test is an appropriate tool.

[43] When reducing data to sessions with the experimental sequence median-mean, the p-value rises to 0.0519 which is not statistically significant any more. See Section 19.6 for sequencing effects.

19.2.2 Influence of information on Nash play

In this section, we want to find out, whether information has a significant influence on Nash play. As test variable, we calculate the percentage of decisions which are optimal in the sense of straight Nash equilibria in each player group and for each rule. Recall, that under the mean rule there exists only one Nash equilibrium which is also straight. Again, we apply a paired t-test, as the standard deviations are similar.

Hypothesis H 8. The average percentage of Nash play per group is smaller (or equal) under full information than under no information under both rules.

We can clearly reject this hypothesis for the mean rule when there are no manipulation costs (p=0.0001). Surprisingly, the contrary holds for manipulation costs, i.e. there is more Nash play under no information. One explanation is the sample size of eight, which makes the result not statistically significant.
When there are no costs, the difference of Nash play is not significant under the median rule (p=0.4729)[44]. When there are manipulation costs, we can reject the hypothesis for manipulation costs under the median rule (p=0.0013).

19.3 Influence of "cost"

In the following, we test whether the variable "cost", i.e. the introduction of manipulation costs has an influence on the frequency of truth-telling. In contrast to the within-subject treatment variables "rule" and "info", the variable "cost" is a between-subject variable. Therefore, we have different numbers of observations and the application of a t-test is no longer sustainable. In this section, we use the Mann Whitney U-Test (see Mann and Whitney (1947)) to rank the observations and test on differences.

Hypothesis H 9. The variable "cost" does not affect the average absolute group deviation in each of the "info" and "rule" treatments.

This hypothesis has to be checked and explained in each of the four treatments as we have seen in Figure 18.10. When considering full information, the average group deviation is significantly smaller when there are manipulation costs. The corresponding p-values are 0.0003 (mean rule) and 0.0167 (median rule)[45]. When no information is available, the difference is not significant for the mean rule (p=0.3958)

[44]When considering only sessions with sequence median-mean, the p-value equals 0.2743.

[45]The p-value of the median rule changes to p=0.0473 when taking into account only sessions with median-mean sequence.

and the median rule (p=0.0525)[46]. This result is compatible to the results we have obtained in Section 19.2.

Although the result is not significant for the mean rule under no information, Figure 18.12(a) suggest a higher rate of truth-telling when there are costs. So, we check the exact truth-telling percentages in another Mann-Whithney-U-Test.

Hypothesis H 10. The variable “cost” does not affect the percentage of truth-telling per group in each of the “info” and “rule” treatments.

These hypotheses can be rejected with high significance. The corresponding p-values are 0.0000 (mean, no info), 0.0000 (mean, full info), 0.0015 (median, no info) and 0.0017 (median, full info).

Hypothesis H 11. The variable “cost” does not affect the percentage of Nash play per group in each of the “info” and “rule” treatments.

These hypotheses correspond to Figures 18.14 and 18.15. We reject them for each of the treatments with p-values 0.0003 (mean, no info), 0.0000 (mean, full info), 0.0000 (median, no info) and 0.0043 (median, full info).

Hence, the influence of manipulation costs is highly significant.

19.4 Influence of “framing”

In our experiment we distinguished between four different framings as explained above. We found that there is no significant difference in the average absolute aggregated value, the percentage of truth-telling or the percentage of Nash play.

19.5 Influence of the rank

In this section we focus on the rank, i.e. the position of the peak in comparison to the other peaks (1= smallest peak, ... , 5= highest peak). As there are five different ranks, a normal t-test is not applicable, so we run an analysis of variances (ANOVA). For each individual we determine the average absolute deviation from the peak in each of the treatments rule/info depending on the rank. Then we tested whether there are differences between the different ranks in an ANOVA and run a post-hoc Scheffe test. We face the problem that an ANOVA is not applicable in all

[46]When considering only sessions with the sequence median-mean, the p-value rises to 0.2320

cases, as the assumption of equal variances has to be rejected in most cases.[47] So, the applicability of an ANOVA is doubtful, but on the other hand, we argue that the difference in variances already leads to the result: Absolute deviation depends on the rank. When ignoring heteroscedasticity, which has been done in several statistical analyses, we found that the mean values differ in all treatments but the median treatment with costs.

The influence of the rank is also represented in the total payoff with respect to the player number. There is no significant difference in the payoffs between the treatments no cost vs. manipulation costs (p-value=0.2627). But there is a significant difference between the player groups when there are no manipulation costs which can be seen in Table 19.3. The p-value of the corresponding ANOVA is 0.000 (degrees of freedom=4, $F = 28.81$). Participants with player number 1,2 and 3 earn significantly more than participants with player number 4 and 5.

	player number					
cost	1	2	3	4	5	Total
no costs	14.69	16.02	15.51	12.94	12.28	14.29
costs	14.08	14.87	14.34	12.78	13.07	13.83
Total	14.59	15.83	15.31	12.91	12.41	14.21

Table 19.3: Average payoff in Euro depending on the player number

19.6 Sequencing effects

During our experiment, we vary the sequence of the mean and median rule. In this section, we test whether the sequence has an effect on the average absolute deviation in the groups. Therefore we run a Mann-Whitney-U-Test for each of the eight combinations of “rule”, “info” and “cost”.

Hypothesis H 12. There is no significant difference in average absolute deviation if the sequence of median and mean is changed in the experiment.

Without manipulation costs, we find no significant differences under the mean rule (no info and full info) and under the median rule with full information. Under the median rule with no information, we have to reject the hypothesis (p-value = 0.0054). This implies that the decisions of persons who played the median rule after the mean rule should be taken in consideration of these circumstances and cannot be taken as history independent. Unless stated otherwise, the analysis of the previous

[47] We ran a Bartlett’s test suggesting the rejection of the assumption of equal variances in all treatments without costs and the treatments under the median rule with manipulation costs.

subsections is not influenced by this result: All significant results remain statistically significant when dropping the data affected by sequencing.[48]

In case of repeating the experiment, median rule experiments should be run before the mean rule. On the other hand, when testing the values of average truth-telling and average Nash play, there were no significant differences between both sequences in all treatments.

As the sample size for the treatment "costs" is very small, we state only for completeness that we find no significant differences in all four combinations of rule and info.

[48] All t-tests were run with both data sets. The p-values changed in a minimal way. Exact values are available upon request.

19.7 Regression

In the previous sections we run parametric and non-parametric tests to analyze effects of single variables. In this section, we summarize these results in a linear regression model. We distinguish between individual and group variables as dependent variable.

19.7.1 A basic regression model with individual variables

In this regression analysis, the dependent variable is **absolute deviation**. As it depends on the individual, we are able to include demographic variables in the regression model.

As in the sections before, the independent variables are **rule** (0 = mean, 1 = median), **info** (0 = no info, 1 = full info), **cost** (0 = no costs, 1 = costs), **framing** (0 = abstract, 1 = firm, 2 = jury, 3 = bank), **rank** (1 to 5), **period** (1-5) and **sequence** (0= median-mean, 1=mean-median). We control for demographic data such as gender (0 = male, 1= female) and econstud (1= studying something related with economics, 0 otherwise). The variables QuestMedian (1= correct median answer, 0 otherwise) and QuestMean (1 = correct mean answer, 0 otherwise) indicate whether the participants correctly answered the questions about which rule can be manipulated. The regression is clustered in the groups of five persons who played together during the entire experiment.[49] Table 19.4 lists the corresponding coefficients.

The stars indicate the level of significance. Rule, info, cost and period are highly significant as well as differences in some ranks. Sequence, gender and QuestMedian play a minor but still significant role in the model. The negative signs of the coefficients indicate that absolute deviation is smaller under the median rule, when there are costs, if the participant is the median voter, if he had understood that manipulation is not possible under the median rule or if she is female. Neither the framing nor the field of study has a significant influence. On the other hand, absolute deviation increases if information is available, if the mean rule is played first, or if the peak is larger than the median peak. According to the regression model, absolute deviation increases each period.

[49]Standard Error is adjusted for 47 clusters in groups of five.

Dependent: abs. dev	Coefficient	(Std. Err.)
rule	-11.231***	(0.803)
info	3.439***	(0.410)
cost	-3.097**	(1.055)
framing 0	0.000	(0.000)
framing 1	0.080	(1.053)
framing 2	-0.673	(0.871)
framing 3	-1.442	(0.736)
rank 1	0.000	(0.000)
rank 2	0.158	(0.589)
rank 3	-5.617***	(0.766)
rank 4	2.832**	(0.872)
rank 5	6.165***	(0.633)
period	1.899***	(0.116)
sequence	2.127*	(0.913)
gender	-2.068*	(0.789)
econstud	0.977	(1.152)
QuestMedian	-2.889*	(1.101)
QuestMean	-0.166	(1.003)
constant	16.659***	(1.149)
N	11280	
R^2	0.168	
$F_{(16,46)}$	110.567	

Significance levels : $*$: 5% $**$: 1% $***$: 0.1%

Table 19.4: Regression analysis

19.7.2 Regression with group variables

In this subsection we consider group variables as the dependent variable. The three different aggregated variables are **group abs dev**, **group Nash** and **group truth**. Each variable was observed four times per group in the four combinations of rule and info. Therefore, we have $4 \cdot 47$ (number of groups) = 188 independent observations.

The dependent variable **group abs dev** indicates the average absolute group deviation over different periods and peak distributions, **group Nash** and **group truth** represent the percentage of Nash play and truth-telling, respectively.

Dependent:	**group abs dev**	**group Nash**	**group truth**
info	0.582	0.025**	0.021**
	(0.390)	(0.009)	(0.007)
rule	-12.980***	0.383***	0.411***
	(0.599)	(0.015)	(0.018)
cost	-3.886***	-0.259***	0.173***
	(0.997)	(0.028)	(0.030)
framing	-0.333	-0.012	0.001
	(0.236)	(0.008)	(0.008)
sequence	0.965	-0.012	0.005
	(0.602)	(0.019)	(0.018)
constant	23.362***	0.510***	0.086***
	(0.725)	(0.024)	(0.018)
N	188	188	188
R^2	0.770	0.836	0.828
$F_{(4,46)}$	108.24	246.12	142.34

Robust standard errors in parentheses

Significance levels : $*$: 5% $**$: 1% $***$: 0.1%

Table 19.5: Group regressions

Table 19.5 shows the regression table. As in the section before, rule and cost are highly significant. The most conspicuous effect we observe is that information is significant for the percentage of Nash play and truth-telling, but not for absolute deviation.

20 Conclusion and outlook

The present text gave an overview over two different voting rules. While the theoretical analysis clearly predicted truth-telling under the median rule, we observed strong deviation in the experiment. As expected, manipulation was higher under the mean rule than under the median rule. By introducing manipulation costs, truth-telling increased under both rules. In the theoretical part, we saw that calculating the Nash equilibrium under the mean rule can become very complex in the presence of manipulation costs in case of deviating from the peak. In contrast, the Nash equilibrium under the median rule becomes unique in the presence of costs. The concept of these different types of equilibria was explained and discussed in detail. We found that in the experiments inefficient equilibria arose extremely rarely.

Beside the strategies we analyzed in the previous sections, we also observed different behaviors such as voting for the group average or waiting for the last round. Especially participants with an extreme rank seemed to suffer from envy or boredom: They sometimes announced values which seem to be totally irrational. Possible explanations are that they just want to do "something" or to harm other players.

We observed that the sequencing of mean and median rule plays a minor role. In the regression analysis, it turned out that gender had a significant effect on manipulation. Women did not deviate as much as men did. It would be interesting to test the strength of this effect in a follow-up study.

What is the interpretation and justification of manipulation costs? At first, they seem implausible. *If manipulation is observable, this implies that true preferences are observable. So why vote?* From the theoretical point of view, costs are a very nice instrument to reduce an infinite set of Nash equilibria to a unique one. On the other hand, they can be interpreted as the expected costs of being caught by a supervising institution in case of deviating. This, of course, only makes sense, if the underlying set of alternatives is somehow objectively observable and not an individually preferred good. For instance, in the framing of a juror evaluating a sport competition this interpretation is very clear. Here, a catalogue of judgement criteria exists and whenever a juror manipulates, there are institutions to check his decisions, give him a penalty and remove him from office. Another interpretation of manipulation costs are psychological overcoming costs for lying. Many individuals feel uncomfortable when they do not tell the truth. For them, the strength of their lie does not play a crucial role. Hence, quasi-fixed manipulation costs – as used in the experiments – are a good way to model these preferences for honesty.

The treatment variable “info” is the one with the most ambiguous interpretations. We gave some explanations in the text. According to theory, it is sufficient for the calculation of the Nash equilibria to know the own rank in a mean rule treatment without manipulation costs. Therefore, a study with a different information structure, namely the indication of the rank, would be very interesting.

A more challenging modification would be a combination of the different chapters of this thesis. It remains an open question how voters with generalized single-peaked preferences would behave in a laboratory experiment if they were confronted with our strategy-proof social choice rule.

Appendix

A The German Federal Parliament and some historical remarks

The political party spectrum in Germany is broad: Not only two big parties exist but several others are important in the federal and state parliaments. The way parties are arranged in the parliament, i.e. the left-to-right-order, has its origin in the French National Assembly (*"Assemblée nationale"*) in 1789 where the left wing represented the progressive socialists and the right wing consisted of the conservative nobility. In Germany this seating arrangement came up for the first time in the Frankfurt National Assembly (*"Frankfurter Nationalversammlung"*) in 1848. Since then, the basic progressive-conservative order remained stable even though lots of new parties have appeared and others disappeared over the years.

Five parties were represented in the 17th German Federal Parliament (2009 – 2013), *"17. Deutscher Bundestag"*. In that legislative period, the governmental coalition consisted of CDU/CSU and FDP. By size, the parliamentary parties were:

- U: Christian Democratic Union (CDU) and Christian Social Union of Bavaria (CSU), total 239 seats
- S: Social Democratic Party (SPD), 146 seats
- F: Free Democratic Party (FDP), 93 seats
- L: The Left (Die Linke), 76 seats
- G: Alliance '90/ The Greens (Die Grünen), 68 seats

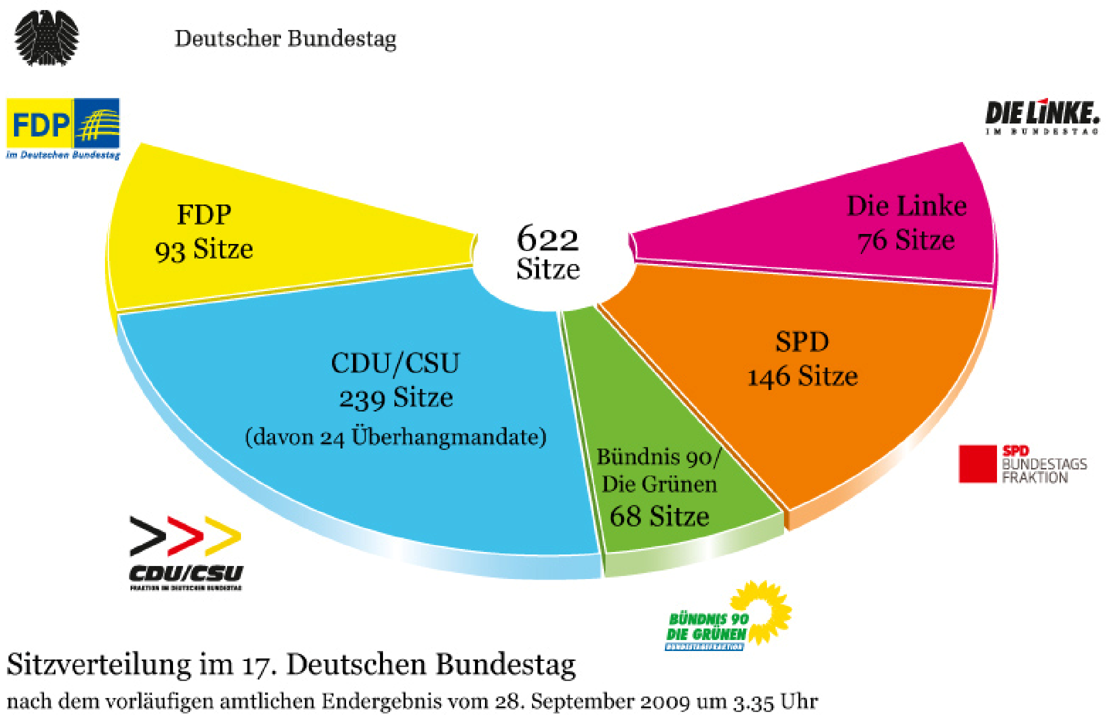


Figure A.1: The German Federal Parliament with seating arrangement *LSGUF*

B Election Results

The results (in percentages) from our survey (S_1, S_2) compared to the second vote of Parliamentary elections at national level (elec) and the corresponding results in the district Karlsruhe (elec KA) are listed in Table B.1

	Union	SPD	FDP	Green	Left
S_1	45.5	25.9	16.2	11.3	1.1
S_2	52.4	18.4	14.4	13.2	1.6
elec 2009	33.8	23.0	14.6	10.7	11.9
elec 2013	41.5	25.7	4.8	8.4	8.6
elec KA 2009	28.6	20.6	17.4	18.4	8.3
elec KA 2013	37.5	22.7	6.0	15.0	6.0

Table B.1: Election Results

C Combinatorial remarks

Calculation of the number 16: The peak divides the ordered set of alternatives into two subsets with k and $(m-1-k)$ alternatives respectively. As we assume the preference ranking to be single-peaked, we know the ranking of the alternatives within these subsets. Combining both subsets we obtain that there are $m-1$ positions where k alternatives can be placed. Summing up over all peaks we get

$$\sum_{k=1}^{m} \binom{m-1}{k-1} = \sum_{\widehat{k}=0}^{m-1} \binom{m-1}{\widehat{k}} = 2^{m-1}$$

which equals 16 for $m = 5$. (Proof see Escoffier et al. (2008))

Example 13. Given the ordering ABCDE and the peak A (i.e. $k = 1$), it is clear that $A \succ B \succ C \succ D \succ E$. For $k = 2$ the peak is B. So we know that $B \succ C \succ D \succ E$. Alternatively, A can be placed in 4 different positions: behind B, C, D or E, i.e. B**A**CDE, BC**A**DE, BCD**A**E and BCDE**A**.

All possible rankings are listed below. In Section 9.1, the set of these 16 rankings is called the **reference set**. For notational convenience we leave the preference symbol ($\succ$) out.
Peak A ($k = 1$): ABCDE
Peak B ($k = 2$): BACDE BCADE BCDAE BCDEA
Peak C ($k = 3$): CBADE CBDAE CBDEA CDBAE CDBEA CDEBA
Peak D ($k = 4$): DCBAE DCBEA DCEBA DECBA
Peak E ($k = 5$): EDCBA

Note that there is a difference between searching for single-peaked preferences on a given party-order and searching for party-orders where a given preference-order is still single-peaked. This is illustrated in the following example: Given the party-ordering CABDE the preference-order ABCDE is single-peaked (see Figure C.1). However, the preference-order CABDE is not single-peaked given the party-order ABCDE (see Figure C.2).

This remark is important for the symmetry structure: Given a preference ranking which is single-peaked with respect to a certain party-ordering, the ranking remains

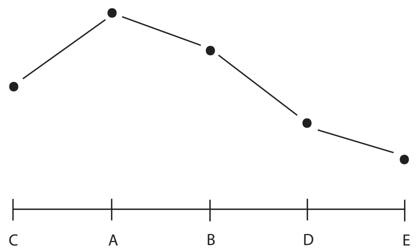


Figure C.1: CABDE

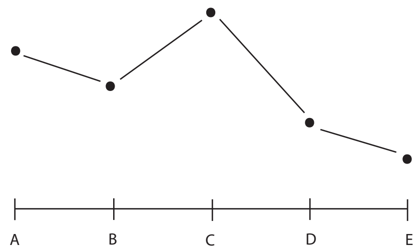


Figure C.2: ABCDE

single-peaked when reversing the order of parties, i.e. mirroring left and right side. Generally, reversing the preference ranking (i.e. the worst alternative is the new favorite) is not single-peaked any more.

D The Questionnaire

The following questionnaire was generated on the Swiss Internet platform `www.onlineumfragen.com`. The entire text is written in German, a translation is available from the author upon request .

OnlineUmfragen.com

Vorwort

Umfrage zu Präferenzen über Regierungskoalitionen

Liebe Studierende,
sehr geehrte Damen und Herren,

danke für Ihre Teilnahme an dieser Befragung. Im Folgenden wird Ihnen eine Frage zu Ihren Präferenzen über Regierungskoalitionen gestellt. Bitte nehmen Sie die Frage ernst, Sie können damit einen erheblichen Beitrag zur Forschung am Institut für Wirtschaftstheorie und Statistik leisten. Vielen Dank für Ihre Mithilfe!

Autor

Veronica Block, Institut für Wirtschaftstheorie und Statistik, Universität Karlsruhe

Zur ersten Frage!

OnlineUmfragen.com

Frage zu Regierungskoalitionen

Stellen Sie sich vor, Sie könnten die Regierung im Bundestag nicht durch Wählen einer Partei, sondern direkt durch Bestimmen einer Regierungskoalition festlegen. Ordnen Sie die links aufgelisteten Koalitionen durch Ziehen auf die rechte Seite so an, dass oben die Koalition steht, die Ihnen als Regierungskoalition am liebsten wäre und unten diejenige, die Sie am wenigsten wünschen. Bitte ordnen Sie dafür **alle** Koalitionen.

Für die bessere Lesbarkeit gilt:
Union steht für CDU/CSU.
LINKE steht für DIE LINKE. ,
GRÜNE steht für Bündnis '90/ Die Grünen.

Bitte begründen Sie Ihre Entscheidung kurz in dem Kommentarfeld unten.

Erstellen Sie bitte eine Rangliste mit 31 Elementen.

Hier nehmen!		*Ihre Rangliste*
Union		1.
SPD	⇄	2.
FDP		3.
LINKE		(...)
GRÜNE		
Union, SPD (aktuelle Regierung)		
Union, FDP		
Union, LINKE		
Union, GRÜNE		
SPD, FDP		
SPD, LINKE		
SPD, GRÜNE		
FDP, LINKE		
FDP, GRÜNE		
LINKE, GRÜNE		
Union, SPD, FDP		
Union, SPD, LINKE		
Union, SPD, GRÜNE		
Union, FDP, LINKE		
Union, FDP, GRÜNE		
Union, LINKE, GRÜNE		
SPD, FDP, LINKE		
SPD, FDP, GRÜNE		
SPD, LINKE, GRÜNE		
FDP, LINKE, GRÜNE		
SPD, FDP, LINKE, GRÜNE		
Union, FDP, LINKE, GRÜNE		
Union, SPD, LINKE, GRÜNE		
Union, SPD, FDP, GRÜNE		
Union, SPD, FDP, LINKE		
Union, SPD, FDP, LINKE, GRÜNE (All-Koalition)		

Bitte begründen Sie Ihre Entscheidung kurz:

Speichern - nächste Frage!

E Results: Generalized single-peaked preferences

Survey 1 (269)	#GSP	t=0	#GSP	t=1	#GSP	t=2	#GSP	t=3	#GSP	t=4
Ordering GLFSU	0	526	0	1107	4	1671	10	2256	62	2977
Ordering GLFUS	0	614	2	1346	5	1854	34	2500	106	3185
Ordering GLSFU	0	668	1	1274	1	1766	13	2408	76	3057
Ordering GLSUF	0	639	1	1387	9	1973	38	2608	117	3276
Ordering GLUFS	0	636	2	1285	13	1851	48	2513	111	3183
Ordering GLUSF	0	518	1	1205	6	1811	43	2485	108	3219
Ordering GFLSU	0	461	0	881	1	1179	3	1531	18	2092
Ordering GFLUS	0	445	1	893	1	1188	4	1571	8	2094
Ordering GFSLU	0	387	1	918	3	1514	15	2087	59	2765
Ordering GFSUL	0	520	6	1197	28	1906	83	2794	192	3611
Ordering GFULS	1	639	2	1256	7	1794	49	2515	109	3146
Ordering GFUSL	3	757	38	1829	90	2567	157	3231	211	3682
Ordering GSLFU	0	582	1	1024	1	1218	7	1622	22	2141
Ordering GSLUF	0	576	0	1020	2	1265	7	1668	22	2213
Ordering GSFLU	0	436	2	1034	4	1565	14	2119	59	2806
Ordering GSFUL	1	729	7	1391	34	2038	107	2924	197	3634
Ordering GSULF	0	523	0	1218	8	1781	34	2460	116	3224
Ordering GSUFL	3	746	20	1602	66	2412	139	3185	210	3712
Ordering GULFS	0	367	0	815	1	1126	4	1569	10	2109
Ordering GULSF	0	383	0	828	1	1143	7	1582	19	2116
Ordering GUFLS	1	617	1	1236	5	1859	37	2451	97	3055
Ordering GUFSL	4	752	34	1637	78	2375	137	3107	195	3595
Ordering GUSLF	0	473	0	1127	5	1778	26	2371	92	3085
Ordering GUSFL	1	533	14	1320	47	2183	112	2975	186	3598
Ordering LGFSU	0	525	6	1165	22	1863	99	2874	193	3656
Ordering LGFUS	8	769	46	1883	94	2567	146	3127	205	3669
Ordering LGSFU	2	775	13	1497	45	2164	127	3112	206	3723
Ordering LGSUF	8	800	23	1683	77	2543	152	3308	215	3770
Ordering LGUFS	11	783	48	1738	89	2477	147	3135	202	3623
Ordering LGUSF	2	539	15	1350	68	2346	135	3127	203	3666
Ordering LFGSU	0	571	11	1243	31	1857	75	2537	155	3385
Ordering LFGUS	0	467	10	1232	34	2017	101	2902	185	3579
Ordering LFSGU	1	516	7	1145	33	1796	77	2541	147	3358
Ordering LFUGS	1	688	13	1411	44	2125	116	3000	198	3645
Ordering LSGFU	1	768	13	1396	32	1918	83	2718	164	3472
Ordering LSGUF	1	716	12	1452	42	2141	112	3016	183	3595
Ordering LSFGU	1	427	5	969	14	1571	61	2465	154	3392
Ordering LSUGF	1	494	12	1251	47	2136	127	3035	195	3628
Ordering LUGFS	1	420	4	1004	16	1572	47	2415	158	3406
Ordering LUGSF	1	521	6	1174	27	1790	88	2617	157	3388
Ordering LUFGS	1	718	7	1333	30	1870	78	2628	166	3441
Ordering LUSGF	2	588	11	1311	34	1956	98	2770	175	3504
Ordering FGLSU	0	460	1	889	2	1177	7	1576	26	2081
Ordering FGLUS	0	460	2	914	2	1220	7	1620	25	2168
Ordering FGSLU	2	497	6	1063	12	1576	36	2217	80	2820
Ordering FGULS	2	424	2	998	5	1518	16	2121	66	2873
Ordering FLGSU	0	545	2	1173	9	1723	29	2271	65	2823
Ordering FLGUS	0	459	1	1082	2	1679	17	2291	58	2910
Ordering FLSGU	1	505	5	1079	10	1617	28	2188	70	2862
Ordering FLUGS	1	457	1	1031	4	1516	16	2100	64	2847
Ordering FSGLU	2	490	4	1083	7	1619	27	2211	64	2746
Ordering FSLGU	1	389	1	803	1	1143	9	1580	23	2029
Ordering FUGLS	2	641	2	1297	5	1853	25	2421	70	2992
Ordering FULGS	1	579	1	1034	3	1284	8	1696	27	2196
Ordering SGLFU	0	578	1	1033	2	1264	5	1650	17	2114
Ordering SGFLU	0	450	2	961	2	1374	5	1861	33	2577
Ordering SLGFU	1	702	2	1321	5	1768	13	2262	63	2839
Ordering SLFGU	1	428	2	987	3	1482	8	2014	32	2608
Ordering SFGLU	0	384	1	900	2	1429	11	1966	33	2564
Ordering SFLGU	0	375	1	806	1	1138	3	1569	11	2000

Table E.1: *con* (connected coalitions) Dataset S_1

Survey 2 (250)	#GSP	t=0	#GSP	t=1	#GSP	t=2	#GSP	t=3	#GSP	t=4
Ordering GLFSU	0	516	0	1050	4	1564	18	2057	58	2690
Ordering GLFUS	0	628	1	1317	9	1815	46	2378	106	2940
Ordering GLSFU	0	626	0	1175	1	1633	22	2190	63	2765
Ordering GLSUF	0	654	2	1355	9	1909	45	2540	117	3118
Ordering GLUFS	0	601	1	1232	12	1732	44	2286	110	2925
Ordering GLUSF	0	517	0	1112	9	1660	26	2249	102	2985
Ordering GFLSU	0	460	0	862	3	1146	7	1528	18	1977
Ordering GFLUS	0	422	0	843	2	1146	9	1573	20	2064
Ordering GFSLU	0	358	1	865	6	1429	18	1915	55	2534
Ordering GFSUL	0	515	6	1125	25	1814	87	2644	172	3320
Ordering GFULS	0	600	1	1173	8	1721	50	2361	117	2976
Ordering GFUSL	5	754	33	1644	71	2279	135	2945	193	3438
Ordering GSLFU	0	555	1	985	4	1230	10	1615	27	2065
Ordering GSLUF	0	565	1	1036	4	1270	15	1676	32	2203
Ordering GSFLU	0	417	1	916	4	1467	18	1994	56	2612
Ordering GSFUL	1	693	13	1331	34	1934	90	2676	167	3286
Ordering GSULF	0	540	1	1231	12	1784	53	2439	125	3093
Ordering GSUFL	4	744	20	1599	67	2347	138	3046	195	3477
Ordering GULFS	0	357	1	791	2	1099	6	1540	22	2056
Ordering GULSF	0	379	1	822	6	1132	10	1539	28	2061
Ordering GUFLS	0	606	2	1231	6	1728	38	2307	103	2893
Ordering GUFSL	2	707	28	1548	73	2220	128	2827	174	3289
Ordering GUSLF	0	521	0	1159	11	1751	30	2288	89	2892
Ordering GUSFL	0	524	11	1271	38	2031	109	2821	183	3397
Ordering LGFSU	0	503	1	1062	16	1703	75	2519	162	3257
Ordering LGFUS	12	797	41	1680	83	2299	128	2848	171	3254
Ordering LGSFU	1	687	6	1269	25	1885	88	2668	171	3320
Ordering LGSUF	4	758	29	1669	75	2456	146	3131	204	3531
Ordering LGUFS	6	730	39	1610	87	2281	135	2875	180	3299
Ordering LGUSF	1	535	21	1342	52	2114	126	2944	195	3465
Ordering LFGSU	0	554	6	1153	22	1691	64	2388	136	3107
Ordering LFGUS	2	513	9	1197	35	1855	81	2645	158	3268
Ordering LFSGU	0	491	4	1067	21	1614	67	2377	140	3107
Ordering LFUGS	1	673	12	1362	42	1983	109	2780	186	3395
Ordering LSGFU	0	667	3	1202	20	1693	62	2414	144	3191
Ordering LSGUF	0	706	13	1422	41	1999	108	2800	181	3406
Ordering LSFGU	0	416	5	977	20	1528	66	2301	138	3070
Ordering LSUGF	0	524	8	1250	45	2037	115	2878	196	3463
Ordering LUGFS	0	408	6	980	20	1528	55	2267	139	3137
Ordering LUGSF	0	500	12	1143	29	1664	69	2404	148	3159
Ordering LUFGS	0	644	4	1219	26	1718	68	2401	159	3232
Ordering LUSGF	3	587	8	1233	33	1855	89	2602	158	3266
Ordering FGLSU	0	471	1	897	4	1203	13	1577	38	2101
Ordering FGLUS	0	430	0	863	3	1179	11	1617	22	2133
Ordering FGSLU	2	479	5	1013	18	1520	33	2028	75	2640
Ordering FGULS	0	434	2	969	5	1452	20	2061	65	2724
Ordering FLGSU	0	559	1	1163	15	1690	33	2161	64	2678
Ordering FLGUS	0	483	1	1073	2	1595	12	2188	70	2845
Ordering FLSGU	1	514	2	1080	14	1602	43	2146	67	2640
Ordering FLUGS	0	474	1	1005	3	1460	18	2094	82	2811
Ordering FSGLU	0	436	2	954	11	1488	27	1999	65	2601
Ordering FSLGU	0	374	2	810	7	1107	14	1498	34	2003
Ordering FUGLS	1	637	2	1264	3	1741	24	2282	66	2847
Ordering FULGS	1	571	2	1019	5	1275	9	1638	22	2119
Ordering SGLFU	0	549	2	988	4	1245	11	1630	20	2047
Ordering SGFLU	0	427	2	892	3	1277	10	1776	42	2445
Ordering SLGFU	0	623	2	1184	5	1623	16	2102	46	2658
Ordering SLFGU	0	425	2	915	5	1422	15	1917	44	2514
Ordering SFGLU	0	336	2	825	4	1299	9	1790	46	2465
Ordering SFLGU	0	358	2	775	2	1075	7	1465	18	1954

Table E.2: *con* (connected coalitions) Dataset S_2

F The Left: an extreme party

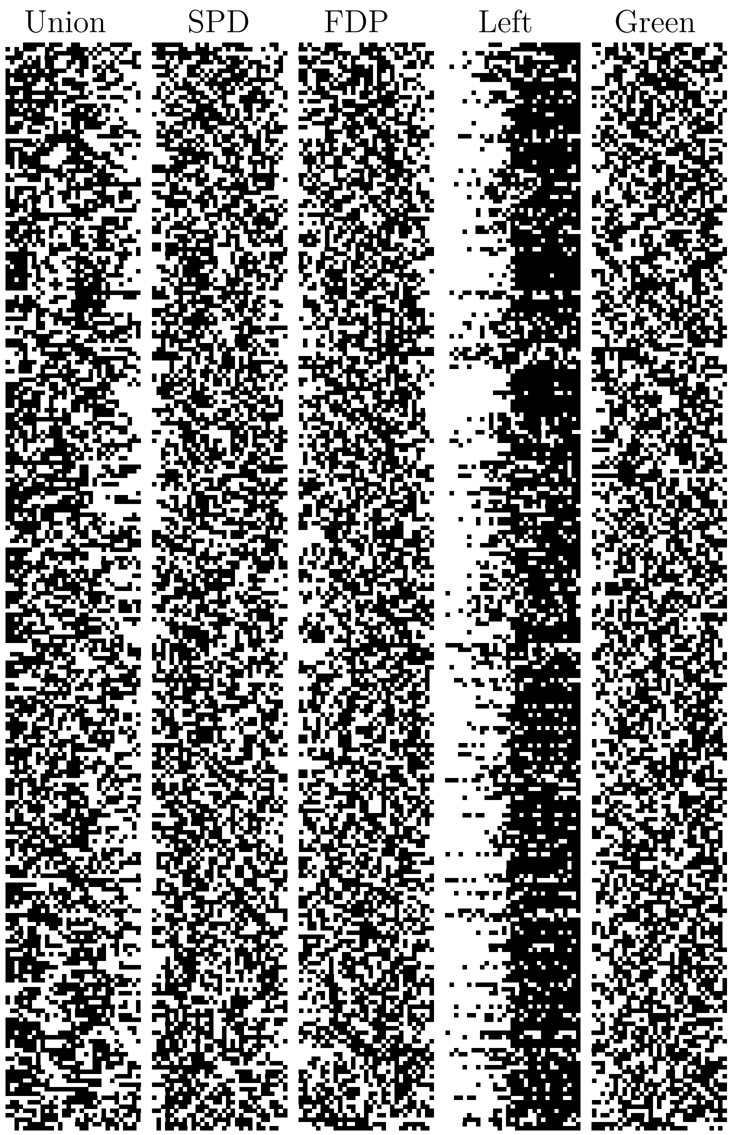


Figure F.1: Special representation for each party: their occurrence in the individual rankings, $n_2 = 250$

G Nash Equilibria

						mean
Peak distr. 1	10	20	50	60	70	42
Equ. no costs	0	0	50	100	100	50
Equ. with costs			–			–
Peaks distr. 2	30	50	60	65	70	55
Equ. no costs	0	0	100	100	100	60
Equ. with costs	30	0	70	100	100	60
Peaks distr. 3	10	20	30	40	50	30
Equ. no costs	0	0	0	100	100	40
Equ. with costs	10	20	0	70	100	40
Peaks distr. 4	10	20	45	60	70	41
Equ. no costs	0	0	25	100	100	45
Equ. with costs	10	20	25	100	70	45

Table G.1: Nash Equilibria for the peaks used in the experiments without and with manipulation costs.

Table G.1 shows the peak distribution used in the experiments with the corresponding Nash equilibria for the mean rule. The equilibrium under the median rule coincides with the peaks. Therefore these values are not listed separately. As we saw in Table 14.3, for the first peak distribution there does not exist an equilibrium with manipulation costs (proof in Section H).

H Proof

Claim: For the peak distribution $(10, 20, 50, 60, 70)$ there does not exist a Nash equilibrium with costs for the special payoff function.

Proof. The proof is by contradictions through a sequence of claims.

Claim 1: $m_1 > 12$

By contradiction, assume $m_1 \leq 12$: The best response of individual 1 lies in $[0, 50]$, so $m \in [0, 20]$ and therefore $m_5 \in [0, 40] \Rightarrow BR_5 = 70$. $\sum_{i\geq 2} x_i \in [0, 50]$ which implies $BR_5 \neq 70$. ↯

Claim 2: $m_1 > 21$

By contradiction, assume $m_1 \in [12, 21.3]$: By the best response function, this leads to $BR_1 = 0$.
Therefore, the mean value is in the interval $m \in [10, 19.3]$. Hence, $m_5 \in [4, 40] \Rightarrow BR_5 = 70$. This implies $m_4 \in [2, 32] \Rightarrow BR_4 = 60$. With $x_5 = 70$ and $x_4 = 60$, we have $m \geq \frac{130}{5} > 19.3$ which is a contradiction. ↯

Claim 1 and Claim 2 lead to $BR_1 = 10 \Rightarrow m \in [21, 82]$.

Claim 3: $m_3 > 33$

By contradiction, assume $m_3 \leq 33$: This leads to $BR_4 = 60$ and $m = m_4$. So, $m \in [21, 33]$. With new calculated $m_3 \in [11, 43], m_5 \in [15, 47]$, we have $BR_3 \geq 50$ and $BR_5 = 70$. So $\sum_i x_i \geq 10 + 50 + 60 + 70 = 190$ and therefore $m \geq 38$. ↯ Contradiction to $m < 33$.

Claim 4: $m_4 \leq 60$

By contradiction, assume $m_4 > 60$: This is equivalent to saying $\sum_{i\neq 4} x_i > 240$. Hence, the mean must be at least $m > 48$. Consider individual 2: $m_2 > 32$, so $BR_2 \in \{0, 20\}$. As $BR_1 = 10$, the sum of the remaining must be $x_2 + x_3 + x_5 > 240 - 10 = 230$. As $x_i \leq 100$ for every i, $x_2 > 30$ which is a contradiction to $BR_2 \in \{0, 20\}$. ↯

Claim 4 implies $BR_4 \geq 60$.

Claim 5: $m_5 \leq 70$

By contradiction, assume $m_5 > 70:$ This implies $m > 56$ and $m_3 > 46, m_2 > 40$. The best response function leads to $BR_3 \leq 70$ and $BR_2 = 20$. So, $\sum_i x_i \leq 10 + 20 + 70 + 60 + 100 = 260 \Rightarrow m \leq 52$. $\lightning$ Contradiction to $m > 56$.

Claim 5 implies $BR_5 \geq 70$ and $m \in [28, 76]$.

Claim 6: $m_2 > 37$

By contradiction, assume $m_2 < 37:$ If $m_2 < 24 \Rightarrow m < 28 \lightning$. If $24 < m_2 < 37$, the best response of individual 2 is $BR_2 = 0$ which implies $m < 33$ and $\sum_i x_i < 165$. As $BR_1 = 10$, the latter one is equivalent to $x_3 + x_4 + x_5 < 155$ such that $x_4 + x_5 \geq 130$. On the other hand, $m < 33$ implies $m_4 < 45$ and $m_5 < 47$ which lead to $BR_4 = 100$ and $BR_5 = 70$. So, $x_4 + x_5 = 170$. $\lightning$ Contradiction to $x_4 + x_5 < 155$.

Claim 6 implies $BR_2 = 20$.

Claim 7: $m_4 \leq 52$

By contradiction, assume $m_4 > 52:$ By Claim 4, $m_4 \leq 60$, this implies that individual 4 determines the mean value and $m = 60$. So, $m_3 \in [50, 70]$ which implies $BR_3 \leq 50$. This leads to $\sum_i x_i \leq 10 + 20 + 50 + 100 + 100 \Rightarrow m \leq 56$. $\lightning$ Contradiction to $m = 60$.

By Claim 3 and Claim 7, $33 < m_4 < 52 \Rightarrow BR_4 = 100$. The updated mean is $m \in [40, 46]$ and therefore $m_5 \in [34, 60] \Rightarrow BR_5 \in \{70, 100\}$. These two values have to be checked in the next claims.

Claim 8: $BR_5 \neq 70$

By contradiction, assume $BR_5 = 70:$ Then, $m = \frac{1}{5}(130 + x_3 + 70) \Rightarrow m_3 = 50 \Rightarrow BR_3 = 50 \Rightarrow m = 50$. $\lightning$ Contradiction to $m \leq 46$.

Claim 9: $BR_5 \neq 100$

By contradiction, assume $BR_5 = 100:$ Then $m = \frac{x_3}{5} + 46 \Rightarrow m_3 = 56 \Rightarrow BR_3 = 20 \Rightarrow m = 50$. $\lightning$ Contradiction to $m \leq 46$.

Hence, no Nash equilibrium exists for the peak distribution $(10, 20, 50, 60, 70)$ with costs and the special payoff function. $\square$

I Utility of best response functions for other peaks

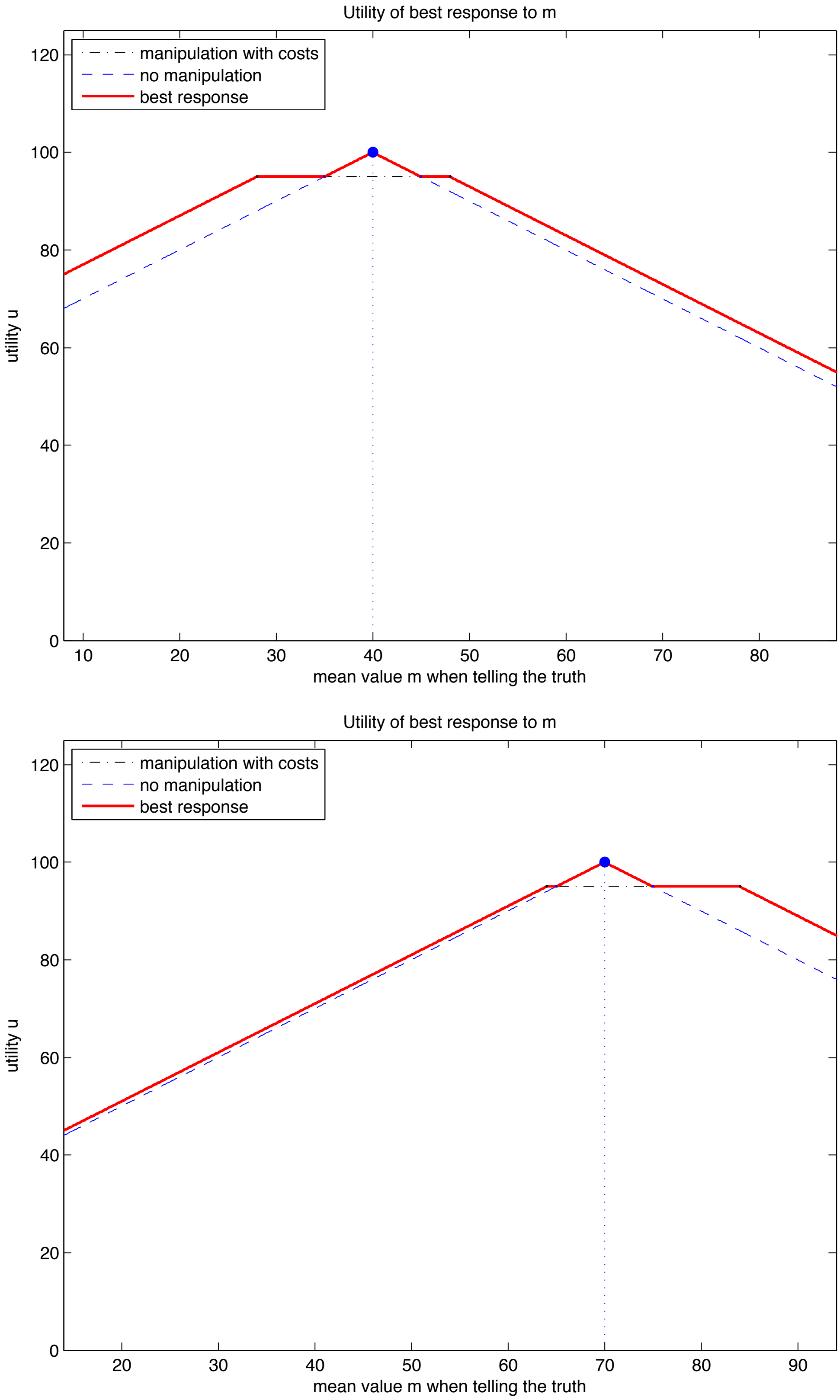


Figure I.1: Utility of best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for linear payoff functions, see also Section 14.3.4

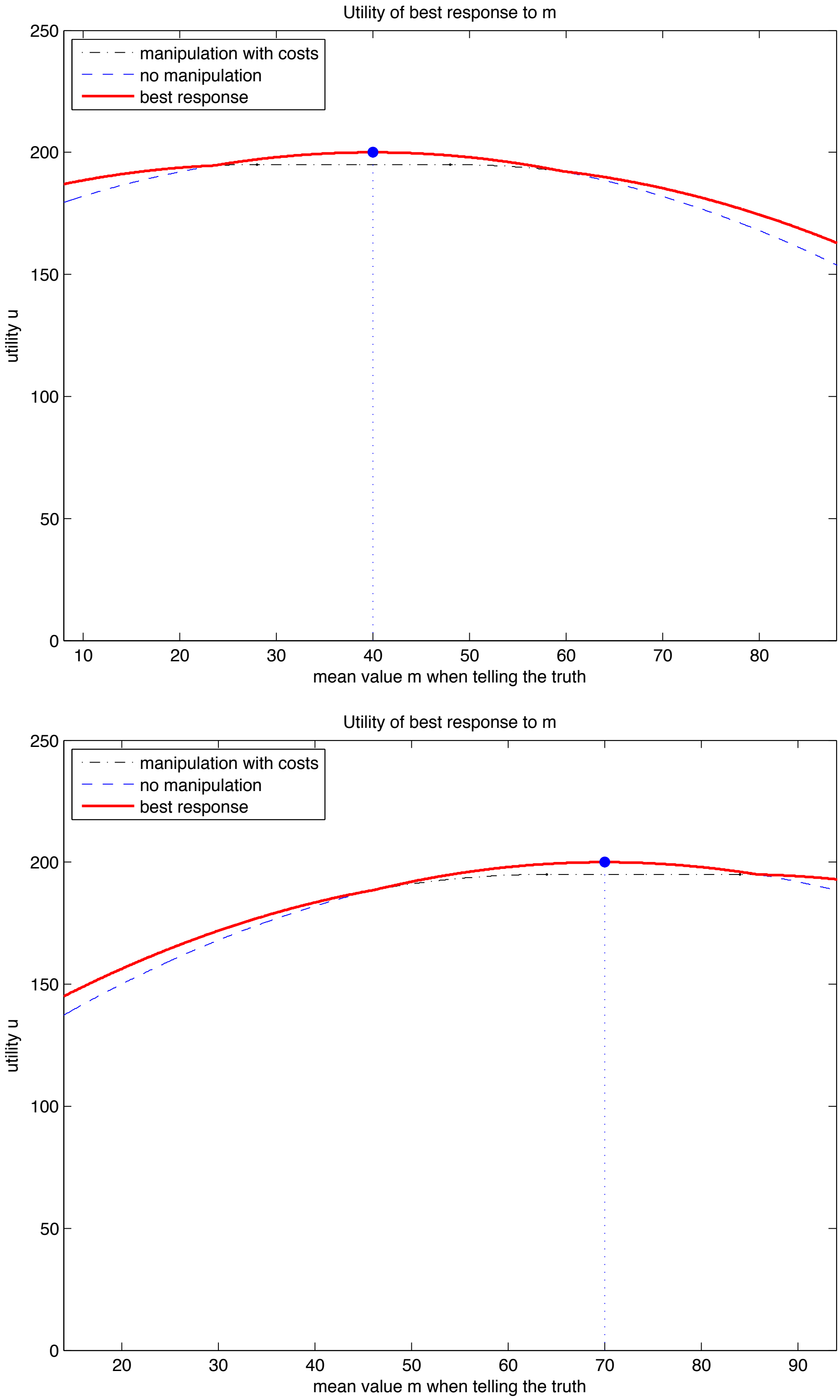


Figure I.2: Utility of best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for quadratic payoff functions, see also Section 14.3.5

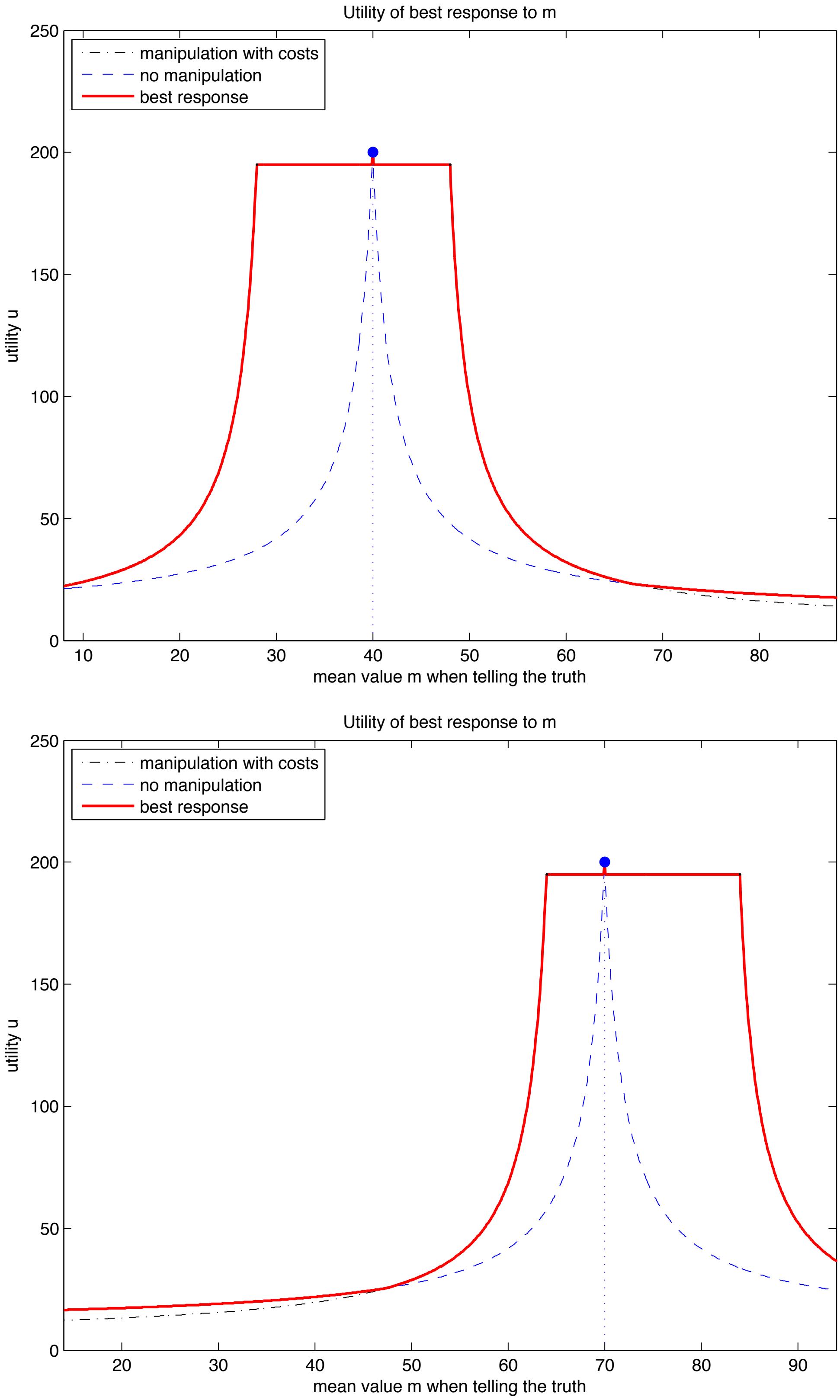


Figure I.3: Utility of best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for special payoff functions, see also Section 14.3.6

J Best response functions for other peaks

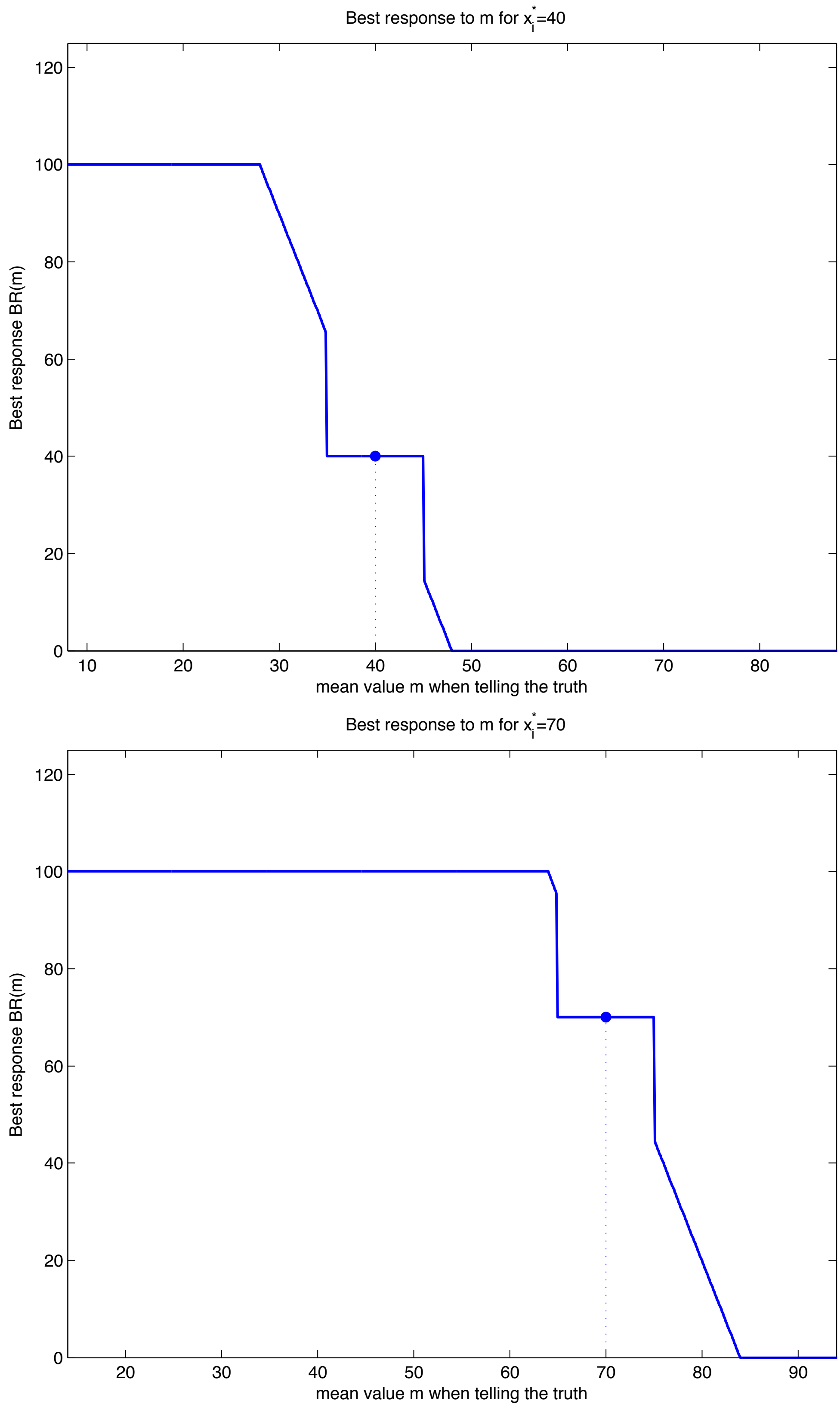


Figure J.1: Best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for linear payoff functions, see also Section 14.3.4

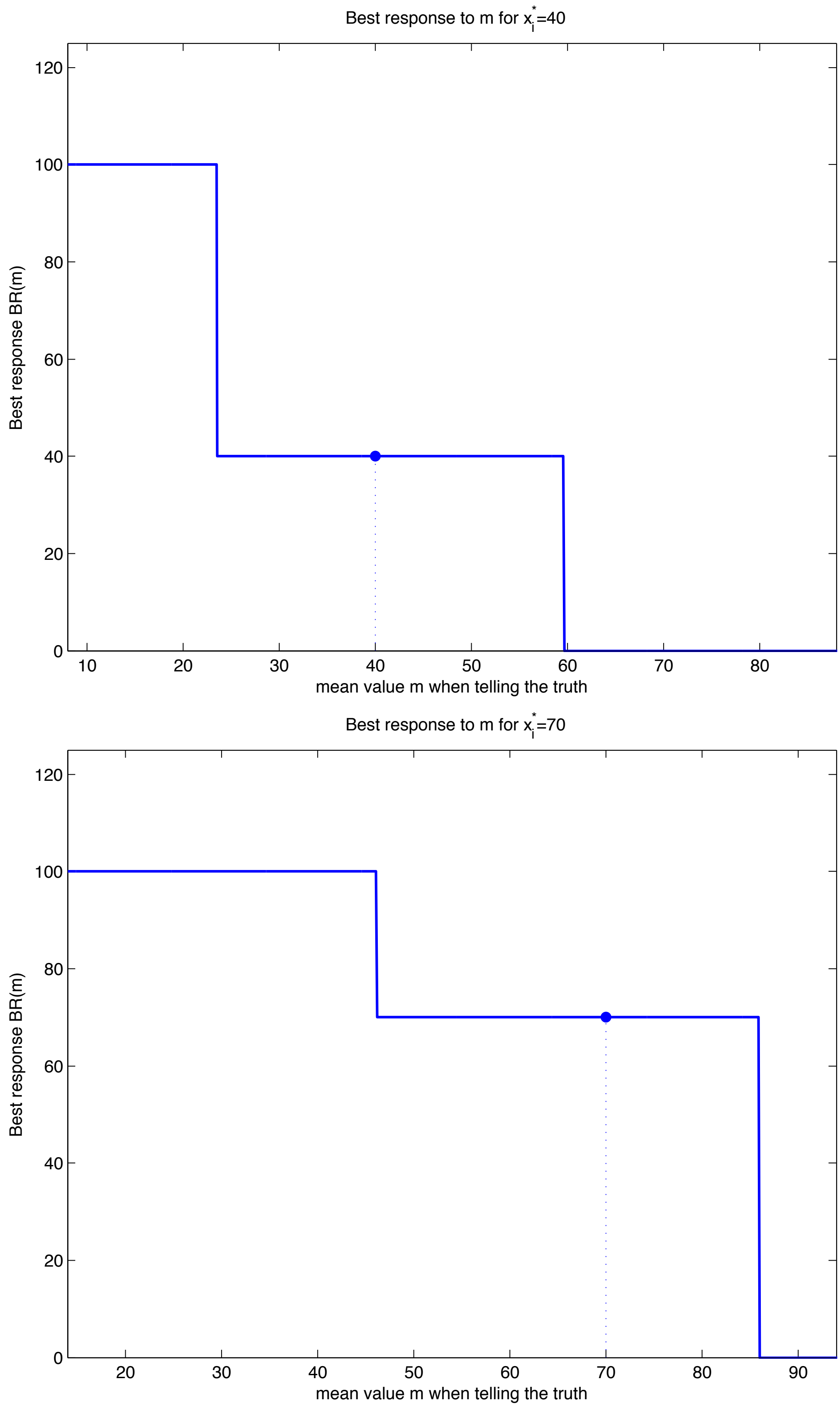


Figure J.2: Best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for quadratic payoff functions, see also Section 14.3.5

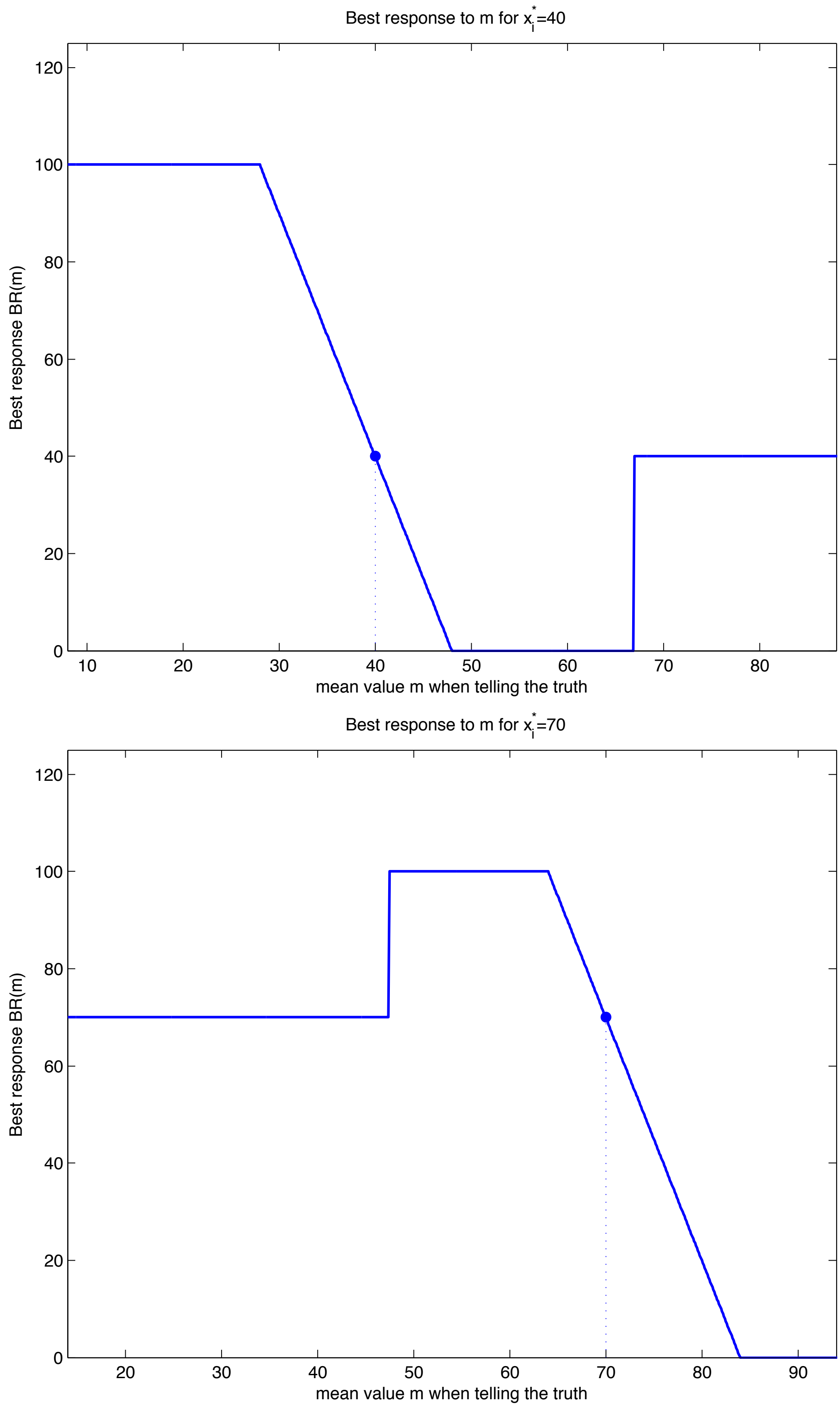


Figure J.3: Best response to m for $x_i^* = 40$ (top) and $x_i^* = 70$ (bottom) for the special payoff function, see also Section 14.3.6

K Handout: Mathematical Introduction

Einführung

Guten Tag. Vielen Dank, dass Sie Zeit gefunden haben, an der Studie teilzunehmen. Wir möchten Sie zunächst mit einigen mathematischen Grundlagen vertraut machen.

Durchschnitt (arithmetisches Mittel)

Der Durchschnitt (arithmetisches Mittel) ist ein Mittelwert, der als Quotient aus der Summe aller Werte und der Anzahl der Werte definiert ist. Bei fünf Werten lautet die Formel zur Berechnung:

$$x = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$$

Beispiel: Wenn die fünf Zahlen $10, 14, 52, 26, 68$ gegeben sind, so ist der Durchschnitt $x = 34$, weil $\frac{1}{5} \cdot (10 + 14 + 52 + 26 + 68) = \frac{170}{5} = 34$

Median

Der Median einer ungeraden Anzahl von Werten ist die Zahl, welche an der mittleren Stelle steht, wenn man die Werte nach Größe sortiert.
Beispiel: Wenn die fünf Zahlen $10, 14, 52, 26, 68$ gegeben sind, so ist der Median $x = 26$, nämlich die mittlere Zahl von $10, 14, \mathbf{26}, 52, 68$.

Betragsfunktion

Es wird mit $|x|$ der Betrag von x bezeichnet. Dieser ist auch als Absolutwert oder $\mathrm{abs}(x)$ bekannt.
Beispiel: $|-15 + 12| = |-3| = 3$.

Minimumsfunktion

Es wird mit $\min(x, y)$ das Minimum von x und y bezeichnet.
Beispiel: $\min(15, |-10|) = \min(15, 10) = 10$.

Veranschaulichung der Auszahlungsfunktion $f_w(x)$

Die Funktion $f_w(x)$ wird im Folgenden mit verschiedenen Werten w verwendet. Sie wird später als Ihre "Auszahlungsfunktion" bezeichnet.

$$f_w(x) = 10 + \min\left(\frac{380}{|x - w + 2|}, \frac{380}{|x - w - 2|}\right)$$

Erläuterung: Beim Wert $x = w$ erreicht die Funktion ihr Maximum und der dazugehörige Wert ist $f_w(w) = 200$. Die Funktion ist achsensymmetrisch zu $x = w$ und ist insbesondere bis zu diesem Wert monoton wachsend und anschließend monoton fallend. Zur Veranschaulichung sehen Sie zwei Beispiele, im oberen Bild ist $w = 12$, im unteren $w = 70$.

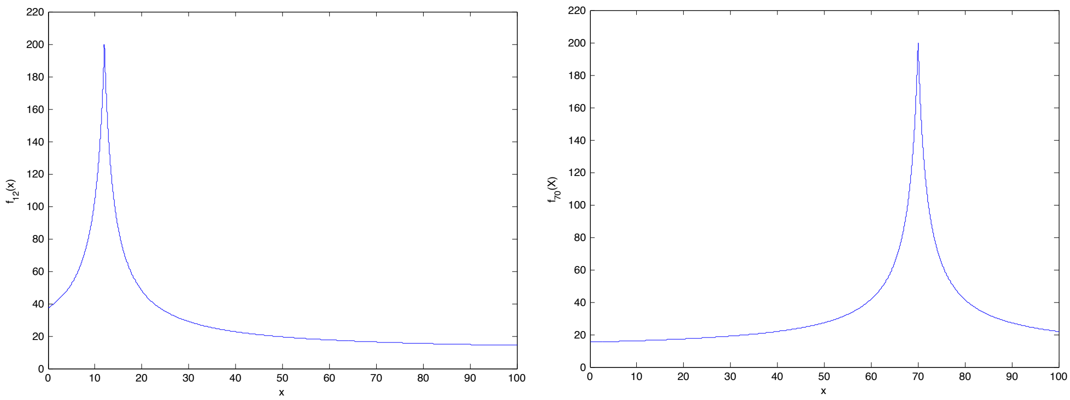


Figure K.1: Zwei Auszahlungsfunktionen, oben: $f_{12}(x)$, unten: $f_{70}(x)$

Bitte beantworten Sie nun einige Verständnisfragen am Bildschirm.

Papier und Stift für Notizen ist bereitgelegt, wir bitten Sie, diese beim Verlassen des Raumes am Platz liegen zu lassen. Außerdem können Sie während der gesamten Studie den Taschenrechner am rechten unteren Bildschirmrand verwenden.

L Handout for the abstract framing without manipulation costs

Anleitung

Sie sind Teilnehmer einer Abstimmung bei der der Wert x bestimmt wird. Hierfür wird Ihnen ein Wert w zwischen 0 und 100 zugeteilt. An der Abstimmung nehmen noch weitere vier Personen teil. Jeder der fünf Teilnehmer macht einen Vorschlag, das heißt er nennt eine Zahl y zwischen 0 und 100. Der Wert x wird nun anhand aller fünf genannten Vorschläge bestimmt. Dazu werden die Vorschläge, also fünf Zahlen zwischen 0 und 100, entweder aufsummiert und durch fünf geteilt (das entspricht dem Durchschnitt aller Vorschläge) oder aber der Größe nach sortiert und der drittgrößte Vorschlag wird gewählt (das entspricht dem Medianverfahren). Sie erfahren jeweils vor der Abstimmung, welches Verfahren benutzt wird. Manchmal werden Sie darüber informiert, welches die zugeteilten Werte der anderen Teilnehmer sind.

Je größer die Differenz zwischen x und w, desto kleiner ist Ihr Gewinn. Ihre Auszahlung ist

$$f_w(x) = 10 + \min\left(\frac{380}{|x - w + 2|}, \frac{380}{|x - w - 2|}\right)$$

Die Auszahlung wird in der Einheit *ECU* angegeben.
100 *ECU* entsprechen 0,60 €.

Die Abstimmung findet in mehreren Runden statt. Der Ablauf jeder Runde ist wie folgt:

1. Sie erfahren das Verfahren der Abstimmung. (Durchschnitt oder Median)
2. Sie erfahren Ihren Wert w und manchmal die Werte der anderen Teilnehmer.
3. Sie machen einen Vorschlag y.
4. Sie erfahren den aus den Vorschlägen berechneten Wert x und Ihre Auszahlung.

M Timetable of the experimental sessions

A = Abstract, F = Company, J = Jury, Z = Bank
Exp 1 and 2: no costs, Exp 3 and 4: costs
Exp 1 and 4: mean - median, Exp 2 and 3: median - mean

Exp	Session	Date	Time	Framing
1	1	04/30/12	09:45	F
1	2	04/30/12	14:00	Z
1	3	04/30/12	16:30	J
1	4	05/02/12	08:00	A
1	5	05/02/12	10:00	J
1	6	05/02/12	17:30	F
1	7	05/03/12	09:00	Z
1	8	05/03/12	11:30	A
1	9	05/03/12	14:00	J
1	10	05/03/12	16:30	A
2	1	06/20/12	14:00	Z
2	2	06/20/12	15:45	F
2	3	06/21/12	11:30	A
2	4	06/21/12	14:00	J
2	5	06/21/12	15:45	Z
2	6	06/22/12	09:00	J
2	7	06/22/12	11:45	A
2	8	06/22/12	14:00	F
2	9	06/25/12	10:00	F
2	10	06/25/12	15:45	Z
3	1	11/23/12	10:00	A
3	2	11/23/12	11:45	A
4	3	11/23/12	14:00	A
4	4	11/23/12	15:45	A

Table M.1: Overview over experimental sessions with date, time and framing

N Truth-telling and Nash play

The following Table N.1 is similar to Table 18.4. For experiments 1 and 2 also those peak distributions are taken into account for which a unique Nash equilibrium exists in the presence of manipulation costs. The percentage of Nash strategy is slightly lower compared to the case when considering all observations.

w/o Distr. 1& 5	truth-telling	Nash play	N°obs
exp 1	0.25	0.63	3800
mean NI	0.09	0.46	1425
mean FI	0.06	0.54	855
med NI	0.46	0.85	950
med FI	0.56	0.85	570
exp 2	0.26	0.64	4000
med NI	0.47	0.90	1000
med FI	0.54	0.89	600
mean NI	0.12	0.46	1500
mean FI	0.09	0.49	900
exp 3	0.39	0.42	800
med NI	0.62		200
med FI	0.69		120
mean NI	0.21	0.31	300
mean FI	0.24	0.19	180
exp 4	0.45	0.47	800
mean NI	0.31	0.32	300
mean FI	0.25	0.31	180
med NI	0.66		200
med FI	0.78		120
Total	**0.28**	**0.60**	**9400**

Table N.1: Overview over truth-telling and Nash strategies in all experiments. Peak distributions without equilibrium with manipulation costs are skipped in all experiments.

Bibliography

Aldrich, J., S. Dorobantu, and M. A. Fernández (2010). The use of the left-right scale in individual's voting decisions. In *APSA 2010 Annual Meeting Paper*.

Ariely, D. (2013). *The Honest Truth about Dishonesty: How We Lie to Everyone–Especially Ourselves*. Harper.

Ballester, M. and G. Haeringer (2007). A Characterization of the Single-peaked Domain. Technical report, Working paper, Universitat Autonoma de Barcelona.

Barbera, S., H. Sonnenschein, and L. Zhou (1991). Voting by committees. *Econometrica: Journal of the Econometric Society 59*(3), 595–609.

Black, D. (1958). *The theory of committees and elections*. Cambridge University Press.

Blais, A., J. H. Aldrich, I. H. Indridason, and R. Levine (2006). Do voters vote for government coalitions? Testing Downs' pessimistic conclusion. *Party Politics 12*(6), 691–705.

Block, V. (2010). Efficient and strategy-proof voting over connected coalitions: A possibility result. *Economics Letters 108*(1), 1–3.

Brams, S., M. Jones, and D. Kilgour (2002). Single-peakednesss and disconnected coalitions. *Journal of Theoretical Politics 14*(3), 359–383.

Cason, T. N., T. Saijo, T. Sjostrom, and T. Yamato (2006, November). Secure implementation experiments: Do strategy-proof mechanisms really work? *Games and Economic Behavior 57*(2), 206–235.

Cherry, T. and S. Kroll (2003). Crashing the party: An experimental investigation of strategic voting in primary elections. *Public Choice 114*(3), 387–420.

Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale.

Coombs, C. (1960). A theory of data. *Psychological review 67*(3), 143–159.

Crawford, V. and N. Iriberri (2007). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica 75*(6), 1721–1770.

Downs, A. (1957). *An economic theory of democracy.* Harper & Brothers.

Dutta, B. and J.-F. Laslier (2010). Costless honesty in voting. Slides of a presentation in Moscow.

Eckel, C. and C. Holt (1989). Strategic voting in agenda-controlled committee experiments. *American Economic Review 79*(4), 763.

Ehlers, L., H. Peters, and T. Storcken (2004). Threshold strategy-proofness: on manipulability in large voting problems. *Games and Economic Behavior 49*, 103–116.

Escoffier, B., J. Lang, and M. Öztürk (2008). Single-peaked consistency and its complexity. In *Proceedings of the 2008 conference on ECAI 2008: 18th European Conference on Artificial Intelligence*, pp. 366–370. IOS Press.

Farrar, C., J. Fishkin, D. Green, C. List, R. Luskin, and E. Levy Paluck (2010). Disaggregating deliberation's effects: an experiment within a deliberative poll. *British journal of political science 40*(2), 333–347.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica 41*(4), 587–601.

Greiner, B. (2004). An online recruitment system for economic experiments. *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, 79–93.

Hirschman, A. (1964). The paternity of an index. *The American Economic Review 54*(5), 761–762.

Kube, S. and C. Puppe (2009). (When and how) do voters try to manipulate? *Public Choice 139*(1), 39–52.

Laakso, M. and R. Taagepera (1979). Effective number of parties: A measure with application to west europe. *Comparative Political Studies 12*(1), 3–27.

List, C., R. C. Luskin, J. S. Fishkin, and I. McLean (2013). Deliberation, single-peakedness, and the possibility of meaningful democracy: evidence from deliberative polls. *The Journal of Politics 75*(01), 80–95.

Mann, H. B. and D. R. Whitney (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics 18*(1), 50–60.

Marchese, C. and M. Montefiori (2005). Mean voting rule and strategical behavior: an experiment. Technical report, Institute of Public Policy and Public Choice-POLIS.

Marchese, C. and M. Montefiori (2011). Strategy versus sincerity in mean voting. *Journal of Economic Psychology 32*(1), 93–102.

Moulin, H. (1980). On strategy-proofness and single peakedness. *Public Choice 35*(4), 437–455.

Nash, J. (1951). Non-cooperative games. *Annals of mathematics 54*(2), 286–295.

Nehring, K. and C. Puppe (2007a). Efficient and strategy-proof voting rules: A characterization. *Games and Economic Behavior 59*(1), 132–153.

Nehring, K. and C. Puppe (2007b). The structure of strategy-proof social choice-Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory 135*(1), 269–305.

Niemi, R. (1969). Majority decision-making with partial unidimensionality. *The American Political Science Review 63*(2), 488–497.

Niemi, R. and J. Wright (1987). Voting cycles and the structure of individual preferences. *Social Choice and Welfare 4*(3), 173–183.

Osborne, M. J., J. S. Rosenthal, and M. A. Turner (2000). Meetings with costly participation. *American Economic Review 90*(4), 927–943.

Radcliff, B. (1993). The structure of voter preferences. *Journal of Politics 55*(3), 714–19.

Renault, R. and A. Trannoy (2005). Protecting minorities through the average voting rule. *Journal of Public Economic Theory 7*(2), 169–199.

Renault, R. and A. Trannoy (2011). Assessing the extent of strategic manipulation: the average vote example. *SERIEs 2*(4), 497–513.

Saijo, T., T. Sjostrom, and T. Yamato (2007). Secure implementation. *Theoretical Economics 2*(3), 203–229.

Satterthwaite, M. A. (1975). Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory 10*(2), 187–217.

Stahl, D. and P. Wilson (1994). Experimental evidence on players' models of other players. *Journal of economic behavior & organization 25*(3), 309–327.

Taagepera, R. and B. Grofman (1981). Effective size and number of components. *Sociological Methods & Research 10*(1), 63.

Tsetlin, I., M. Regenwetter, and B. Grofman (2003). The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare 21*(3), 387–398.

Van Deemen, A. and N. Vergunst (1998). Empirical evidence of paradoxes of voting in Dutch elections. *Public choice 97*(3), 475–490.

Van der Straeten, K., J.-F. Laslier, N. Sauger, and A. Blais (2010). Strategic, sincere, and heuristic voting under four election rules: an experimental study. *Social Choice and Welfare 35*(3), 435–472.

Yamamura, H. and R. Kawasaki (2013). Generalized average rules as stable nash mechanisms to implement generalized median rules. *Social Choice and Welfare 40*(3), 815–832.