Collective Decisions on Conditional Topics

An Empirical Study of the Impact of Nonseparable Preferences

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES
DOCTOR RERUM POLITICARUM

AN DER

FAKULTÄT FÜR WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN
DER RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

vorgelegt von

Andreas Fleig

Heidelberg, Juni 2013

DEAN

Prof. Dr. Jürgen Eichberger, Department of Economics, University of Heidelberg

EXAMINER

Prof. Dr. Daniel Finke, Department of Political Science, University of Heidelberg

Prof. Christoph Vanberg, Ph.D., Department of Economics, University of Heidelberg

Date of Disputation: 04.10.2013

FINANCIAL SUPPORT

The completion of this work was promoted by the Baden-Württemberg Foundation and the Completion grant of the Graduate College of the University of Heidelberg.

Scientific contributions

As part of my research over the last three years the following scientific contributions have evolved:

CONFERENCE PRESENTATIONS

- Fleig, Andreas. 2011a. The Merits of Adding Complexity: Conditional Preferences in Spatial Models of EU Politics. Presented at the 6^{th} ECPR general conference at the University of Iceland, Reykjavik.
- Fleig, Andreas. 2011b. Pooling vs. Delegation: Collective decision-making under uncertainty. Presented at the Annual Meeting of the working group for action and decision theory of the German association for political science (DVPW), University of Kiel.
- Fleig, Andreas. 2012. Motivated by the Process: How decision-making procedures shape the motivation function under majority rule. Presented at the 3^{rd} Thurgau experimental economics meeting (theem) at the University of Konstanz, Kreuzlingen (CH).

ARTICLES

- Finke, Daniel and Andreas Fleig. 2013. "The Merits of Adding Complexity: Non-Separable Preferences in Spatial Models of EU Politics." *Journal of Theoretical Politics* 25:547-576.

WORKING PAPERS

- Finke, Daniel and Andreas Fleig. "Identifying Social Preferences in Majority Decisions." *mimeo*.
- Fleig, Andreas and Daniel Finke. "Delegation, Uncertainty and Social Preferences in Majority Decisions?" *mimeo*.

These contributions are also part of my dissertation. I indicated this in the according chapters and sections by corresponding references.

Abstract

Analytical politics investigates collective decision-making in political systems. Such voting behavior in groups takes place in parliaments, committees or the board of local football clubs. It is a frequent object of study for theoretical as well as empirical analysis. Previous contributions have demonstrated well the stabilizing effect of procedural rules, such as agenda-setting or multi-chamber systems, for collective decisions. These rules are applied in many institutions, such as the European Parliament or the German Bundestag. Their main purpose is to ensure reliable policy.

Previous work continually used the restrictive assumption of separable preferences. This assumption implies that different aspects of a question do not influence each other. The limited validity of this hypothesis is apparent even in everyday situations. For example, the enjoyment of a delicious meal depends on the combination of food and drink. When choosing between fish and venison for dinner you also have to consider the question of which sort of wine to have with the meal; white with fish, and red with venison. This form of interdependence also occurs in legislation. For example, the savings determined in the Greek budget influence the preferences of the German public for financial assistance to Greece.

The assumption of separable preferences is therefore in the critical focus of theoretical research. This literature discusses the impacts of and solutions to nonseparable preferences in detail. The analysis suggests an increased complexity for every decision-making process affected by nonseparable preferences. This complexity leads to difficulties in the operationalization of nonseparable preferences and is one of the reasons that there are too few empirical examinations. In addition, the stabilizing properties of institutional arrangements identified under the assumption of separable preferences are in question. The goal of this study is to close this gap between theory and empiricism.

I investigate nonseparable preferences by conducting a laboratory experiment, which allows comprehensive environmental control. This facilitates the operationalization of nonseparable preferences. First, I prove the relevance of nonseparable preferences for analytical research on social interaction. The experiment is therefore completed by empirical case studies. Next, I investigate the effects of nonseparable preferences on collective and individual decision-making in the laboratory. Finally, I assess my contribution with respect to current research in social science and discuss possibilities to more accurately model of human behavior.

The dissertation is structured as follows. I start in chapter 1 with the presentation of my research question and design. In chapter 2 the concept of nonseparable preferences is further clarified by means of exemplary case studies. It also discusses the theoretical foundations of nonseparable preferences. My hypotheses are specified along common concepts used in the literature. Based on empirical data the relevance of nonseparable preferences for political science research is demonstrated in chapter 3. Next, chapter 4

presents the design of the laboratory experiment. The effects of nonseparable preferences on collective decision-making are examined in chapter 5. Subsequently, determinants for the motivation function of individuals are scrutinized in chapter 6. In chapter 7 I report the results of the post-experiment survey. All findings are evaluated in chapter 8, where I focus on detailing their usefulness to future research on human behavior. Finally, chapter 9 summarizes the study and lists possibilities to further expand research in this area.

Zusammenfassung

Die Analytische Politikwissenschaft befasst sich mit kollektivem Entscheidungsverhalten im Kontext politischer Systeme. Solche Gruppenentscheidungen betreffen Abstimmungen in Parlamenten, Ausschüssen oder dem Vorstand des örtlichen Fussballklubs. Sie sind ein häufiger Untersuchungsgegenstand theoretischer sowie empirischer Analysen anhand derer die stabilisierende Wirkung fester Verfahrensregeln wie Agendasetzung oder eines Mehrkammernsystems für kollektive Entscheidungen mehrfach nachgewiesen wurde. Diese Regelungen finden in zahlreichen Institutionen wie dem Europäischen Parlament oder dem Deutschen Bundestag Verwendung. Ihre zentrale Aufgabe ist die Sicherstellung verlässlicher Politikentscheidungen.

Die bisherigen politikwissenschaftlichen Beiträge verwendeten fortwährend die restriktive Annahme separabler Präferenzen. Diese besagt, dass einzelne Aspekte einer Frage sich nicht gegenseitig beeinflussen. Die limitierte Validität dieser These zeigt sich schon in alltäglichen Situationen. Der Genuss eines guten Essens hängt von der Kombination von Speisen und Getränken ab; die Wahl zwischen Rot- und Weißwein ist nicht von der Entscheidung für ein Fisch- oder Wildgericht zu trennen. Diese Form von Abhängigkeiten tritt auch in Gesetzesvorhaben auf. So beeinflussen die beschlossenen Einsparungen im griechischen Haushalt die Bereitschaft der deutschen Öffentlichkeit zu einer finanziellen Unterstützung Griechenlands.

Die Annahme separabler Präferenzen steht deshalb im kritischen Fokus der theoretischen Forschung, welche eingehend Auswirkungen von und Lösungskonzepte für Nichtseparabilität diskutiert. Diese Analysen weisen auf eine gesteigerte Komplexität im Entscheidungsfindungsprozess hin. Die resultierenden Schwierigkeiten bei der Operationalisierung sind zum einen der Grund für die fehlenden empirischen Untersuchungen. Zum anderen stellen sie die unter der Annahme separabler Präferenzen identifizierten stabilisierenden Eigenschaften institutioneller Regelungen in Frage. Das Ziel meiner Arbeit ist es, die hier herrschende Lücke zwischen Theorie und Empirie zu schließen.

Dies erfolgt anhand eines Laborexperiments, welches durch eine umfassende Modifikation der Rahmenbedingungen die Operationalisierung von Nichtseparabilität ermöglicht. In einem ersten Schritt weise ich die Relevanz des Konzepts der Nichtseparabilität für die analytische Erforschung sozialer Interaktion nach. Hier wird das Experiment mittels empirischer Fallstudien komplettiert. Anschließend untersuche ich die Auswirkungen von Nichtseparabilität auf kollektives und individuelles Entscheidungsverhalten im Labor. Zum Abschluss ordne ich meinen Beitrag in den Kontext aktueller sozialwissenschaftlicher Forschung ein und diskutiere die Forderungen nach einer realistischeren Beschreibung menschlichen Handelns.

Kapitel 1 beginnt mit der Darlegung der Forschungsfrage und des Forschungsdesigns. In Kapitel 2 wird mit Hilfe exemplarischer Fallstudien das Konzept nichtseparabler Präferenzen weiter präzisiert. In diesem Kapitel erläuterte ich außerdem die theoretischen

Grundlagen sowie die konkrete Operationalisierung von Nichtseparabilität und formuliere meine Hypothesen über die Folgen ihrer Nichtberücksichtigung. Anhand empirischer Daten wird in Kapitel 3 die Relevanz von Nichtseparabilität für die politikwissenschaftliche Forschung aufgezeigt. Anschließend stellt Kapitel 4 das Design des Laborexperiments vor. Die Auswirkungen von Nichtseparabilität auf kollektive Entscheidungsfindung werden in Kapitel 5 untersucht. Darauf aufbauend werden die Determinanten der Motivationsfunktion von Individuen in Kapitel 6 analysiert. Kapitel 7 diskutiert die Ergebnisse des post-experimentellen Fragebogens. Die gewonnen Erkenntnisse werden in Kapitel 8 bewertet. Dabei liegt der Fokus auf dem Nutzen für die zukünftige Forschung über menschliches Verhalten. Schließlich fasst Kapitel 9 die Studie zusammen und nennt mögliche Erweiterungen des Forschungsbereichs.

Contents

List of Tables

List of Figures

List of Abbreviations

θ	An actor's unconditional ideal poin
A	Matrix of saliences
a	Single element of the matrix of saliences
ACP	African, Caribbean and Pacific Group of States
AIC	Akaike information criterion, a measure of the relative goodness of fit of a statistical model (Akaike 1973)
ALDE	Alliance of Liberals and Democrats for Europe
AMT	Amazon Mechanical Turk
APM	Accumulated payoff mechanism
BW	German Federal Armed Forces (translation of Deutsche Bundeswehr)
CI	Confidence intervals
CP-nets	Conditional preference networks
CRRA	Constant relative risk aversion
D	Somers' D, an ordinal measure of association introduced by Somers (1962)
d	Metric distance
D^+	A measure for advantageous inequality in the inequality averversion model of Fehr and Schmidt (1999)
D^-	A measure for disadvantageous inequality in the inequality averversion model of Fehr and Schmidt (1999)
DB	Deutsche Bahn
DEU	Decision-making in the European Union
DGP	Data-generating process
DVPW	German association for political science (translation of Deutsche Vereinigung für Politische Wissenschaft)
ECHA	European chemicals agency
ECPR	European consortium for political research
EDE	Experimenter demand effect
EP	European Parliament

EPP-ED	European People's Party-European Democrats
ERC	Equity, reciprocity and competition
ESA	Economic Science Association
ESM	European stability mechanism
FBNE	Feedback Nash equilibrium
Gini	Gini coefficient, a measure of statistical dispersion (Gini, 1912)
HSP	Heterogeneous social preferences
i	Index of actor
IIA	Independence of irrelevant alternatives
iid	Independent and identically distributed
INET	Institute for new economic thinking
IP	Ideal position of an actor in a spatial model
IPO	Initial public offer
MAE	Mean average error
MCMC	Monte Carlo Markov Chain
MEP	Member of the European parliament
MPNE	Markov-perfect Nash equilibrium
N	Number of observations
n	Number
Nash product	Maximand of the Nash Bargaining Solution (Nash, 1950)
NBS	Nash bargaining solution
NSP	Nonseparable preferences
OLNE	Open-loop Nash equilibrium
ORSEE	Online recruitment system for economic experiments
OSLA	One-step look-ahead algorithm
P	Probability
POL	Pooling; a specific voting procedure in my laboratory experiment
QMV	Qualified majority voting
QRE	Quantal response equilibrium
r	Minkowski r-metric
RCV	Roll-call vote

REACH	The European Union regulation on registration, evaluation, authorisation and restriction of chemicals
RMDSM	Random multiple decision selection mechanisms
RRPM	Random round payoff mechanism
RUMM	Random-utility mixture model
S&K	Sauermann and Kaiser (2010), I adopt their general laboratory environment for my experimental design
SCF	Social choice function
SD	Standard deviation
SE	Standard error
SEQ	Sequential delegation; a specific voting procedure in my laboratory experiment
SIM	Simultaneous delegation; a specific voting procedure in my laboratory experiment
SPE	Subgame-perfect equilibria
SQ	Status quo
SQP	Sequential quadratic programming
SSC	Statistical software components
SSS-V	Sensation Seeking Scale V
SVO	Social value orientation
theem	Thurgau experimental economics meeting
tm	Text mining
U	Utility level
U.S.	United States of America
w	Weight of spatial dimension
WED	Weighted Euclidean distance
WEIRD	Western, educated, industrialized, rich and democratic
WTO	World Trade Organization
z-Tree	Zurich Toolbox for Readymade Economic Experiments

1 Introduction

The debate surrounding a budget for a political program is a matter of everyday politics. The preferences of the politicians deciding on the size of the budget depend on the politically determined aim of the program; e.g., who is entitled to a subsidy or who is exempt from a certain regulation. If the purposes of both the program and politics match, one favors the allocation of more money. If they contradict each other, one aims to keep the financial support low. A relationship which is as obvious as it is banal. A current and concrete example of such conditionality is not hard to find: the austerity measures determined in the Greek budget affected the willingness of the German public to offer financial assistance to Greece during the last two years.[1]

Such dependencies are not restricted to a financial dimension. The passing of exploitation rights from a principal to an agent most certainly depends on the compliance of intentions between the two. The principle of subsidiarity is applied more often if the lower administrative level is politically in line with the higher level of administration. Preferences for the distribution of voting rights are connected to the presence of comrades and competitors in the electorate. Formulating the actual question of a referendum (e.g., adding or omitting parts) exerts a lot of influence on the public vote.

When political scientists analytically describe such conditionalities with theoretical models, even simple compounds can very quickly become complicated. Therefore, the individual components are often examined separately, i.e., the result of one decision is investigated without considering the outcome of other decisions. Possible interactions are denied. This simplification entails the risk of obtaining incorrect results and drawing the wrong conclusions.

Analytical politics incorporates conditionalities as nonseparable preferences (henceforth NSP) into its models. The most basic and simple definition for NSP is that the preferred choice of an actor on one issue changes with the result of a second issue: how much one wants to spend on a policy program depends upon exactly to what degree it supports one's own policy-making intentions.

Survey and referendum research relies on the measurement of individual preferences. It is essential to measure them correctly and precisely. The current trend suggests that these fields will become even more important at the local as well at the federal level in the future. Beedham (1993, p. 5) identified a development in the nature of democracy indicating

[1]Source: Public poll by the Forsa survey research institute. The survey was conducted on September 15th and 16th 2011 and the number of respondents was 1002 German citizens (Forsa, 2011).

"a shift from representative democracy to direct democracy",[2] i.e., towards more citizen participation and public voting.[3] It is inevitable that a survey or referendum enables its participants to express their preferences in an undistorted, clear and complete manner. A single binary vote may not be enough if the topic in question contains interrelated issues. This would risk mistaking the respondents' conditional response as their genuine first preference (Lacy, 2001a). Thus, not the genuine response is captured but only the reply adjusted to another circumstance. Conclusions based on such results are inaccurate and misleading.

In my study, I empirically analyze whether and, if it is the case, how the nonseparability of preferences influences decision-making processes. This enables me to identify proceedings which avoid the biased results otherwise obtained. My contribution does not refer to a single empirical case or specific data set. Rather, I draw attention to the phenomenon of NSP and gather contributions from various fields to synthesize a single and consistent assessment. This lays the coherent foundation for research on NSP, which has until now been lacking.

The following sections present an overview of my study. I start by explaining my research question in more detail in Sec. 1.1. Next, I discuss previous research in Sec. 1.2 in order to point out the research gap which I aim to fill. This section also concisely addresses the areas of scientific research which might gain from applying my results. I follow these findings up with a delimitation of the focus of the study in Sec. 1.3 in which I describe my research design and substantiate the use of a laboratory experiment. The last section presents a further outline of the study.

1.1 Research question

Political science deals with a broad variety of issues. These range from everyday decisions to simple judgments and complex problems. In the realm of politics decisions are usually made in a group and not individually.[4] Such collective decision-making takes place in parliaments, committees, or the board of local football clubs.[5] Analytical politics aims to identify the determinants for decision-making by means of a formalized

[2]Morris (1999, p. 23) characterized this development as "shift from representational (Madisonian) to direct (Jeffersonian) democracy" and as the "fundamental paradigm that dominates our politics".

[3]Cf. Dalton et al. (2001) for a summary of feasible institutional designs in the continuum between direct and representative democracy. In general, the question whether political decisions should be taken rather in parliament or by referendum is controversial. For example, proponents of more direct democratic participation repeatedly point out the endemic irresponsibility of much too low estimated costs in government procurement (Merkel, 2010). Yet, opponents insist on the non-negotiable democratic legitimacy of parliament decisions which are more reliable and robust against short-term mood changes and propaganda. Cf. McConnachie (2000) for a discussion of the advantages and disadvantages of a public vote and Fiorino and Ricciuti (2007) on the determinants of direct democracy.

[4]Miller (1997, p. 1181) pointed out that the "mediation of groups distinguishes how individuals participate in political settings from the way that they participate in markets."

[5]In the tradition of Buchanan and Tullock (1962) this means "ordinary politics" which operates within the already established "constitutional stage" (cf. Buchanan, 1986).

approach[6]. This includes multiple criteria such as the wishes and beliefs of individuals (e.g., Dyer et al., 1992; Wallenius et al., 2008) and institutional settings (e.g., Peters, 2000).

Collective decision-making was and is a frequent object of study for theoretical as well as empirical analyses (Kirsch, 2004).[7] Previous contributions have successfully demonstrated the influence and stabilizing effect of procedural rules, such as agenda-setting, majority voting, or multi-chamber systems that structure the legislative process (e.g., Bottom et al., 2000; Miller and Hammond, 1990).[8] These rules are applied in many institutions such as the European Parliament or the German Bundestag. Their central purpose is to ensure reliable policy-making (Tsebelis, 2002) and prevent infinite debates (cf. McKelvey's Chaos Theorem, McKelvey, 1976; Shepsle and Cox, 2007).

Individual and collective decisions are mostly more complex than simple 'yes or no' questions.[9] Imagine the preparation of an exquisite dinner. The decision of which wine is the most appropriate hinges on the main course one has selected. The overall success of the dinner depends on the combination of drink and food: with fish I prefer white wine; if venison is served I opt for red wine. Another example that illustrates this complexity is that of two neighboring villages which both face decisions on development plans of their municipality. The first village plans to use certain areas of farm land to build homes and for industrial purposes. Unfortunately, the second village is planning to build a vast incineration plant on fields adjacent to this neighborhood. Obviously, the latter plan might reduce the value of land if the first village chooses the wrong fields. Again, the preferred choice in one decision depends on the outcome of another.[10]

Highly complex and multidimensional legal matters are typical of the field of politics.[11] Conflicts on legislative proposals can usually be subdivided into multiple, interrelated issues (e.g., von Prittwitz, 2007; Tsebelis, 1994). As a consequence, legislators' preferences on one issue may depend on the (expected) outcomes of other issues (Hinich and Munger, 1997). One well-known case is the conditionality in legislators' preferences, as discussed above, on the characteristics of and budget for a political program. In fact, a "large [amount of] literature in economics demonstrates that people cannot be assumed to have separable preferences on public spending issues" (Lacy and Niou, 2000, p. 8).

Many contributions to research in political science agree that NSP exist (e.g., Brams et al., 1997; Lacy, 2001a). They occur in combinations of different subjects, various fields, vari-

[6]Cf. Morton (1999, Chap. 2) for an excellent comparison of formal as well as non-formal models.

[7]Cf. Kirsch (2004) for a comprehensive overview on the fields of new political economy and institutional analysis. He investigated collective decision-making in political systems in general. More specifically, Schofield (1996) discussed social choice theory in economic and political decisions.

[8]Frohlich and Oppenheimer (2007, p. 363) discussed an alternative solution to reach stability by introducing a "culturally accepted conception of justice within a utility function." Although this seems promising, I will follow the standard response of restricted procedures respectively preferences.

[9]Braumöller (2003) argued that theories which posit complex causation, or multiple causal paths, pervade the study of politics. He offered an extensive selection of examples for various fields of political science.

[10]Sec. 2.2 provides a multitude of further examples.

[11]Krehbiel et al. (2004, p. 251) pointed out that "social choice theoretic results clearly demonstrate the analytical significance of dimensionality."

ant questions, a variety of individual or collective actors, etc. Many theorists have discussed the possible impacts and consequences in detail (e.g., Denzau and Mackay, 1981; Kramer, 1972). Nevertheless, nearly all empirical analyses assume for reasons of simplicity that preferences are separable (e.g., Blundell, 1999; Schneider et al., 2010). This demanding assumption requires that "preferences on every issue and set of issues are independent of - or, can be separated from - the outcomes of other issues" (Lacy, 2001b, p. 240); an assessment that seems rather daring in many situations.

Neglecting conditionalities also outlasts recent improvements. Political scientists have made rapid progress in measuring policy positions, conducting expert interviews (e.g., Benoit et al., 2005; Thomson et al., 2006), computer-aided text analysis (e.g., Laver et al., 2003; King and Hopkins, 2007) or dynamic ideal point estimation (e.g., Jones et al., 2009; Martin and Quinn, 2002). This recent work has increasingly tried to include information on preferential distributions as well as the collection of actor-specific salience[12]. This is a clear improvement compared to focusing solely on an actor's first preference.[13] However, the existing methods build on the assumption of separable preferences. Thus, an important aspect remains unspecified.

I do not argue that every research project should implement nonseparability. But one ought to investigate its possibilities. If one does not take them into account, this aspect of methodology should be addressed. Previous theoretical analyses have identified an increased complexity for every decision-making process affected by NSP (cf. Enelow and Hinich, 1984). This complexity leads to difficulties in model operationalization and is the reason this topic has only seen little empirical examination. Moreover, it questions the stabilizing properties of institutional arrangements identified under the assumption of separable preferences. This has implications even for the most basic principles of institutional analysis.

The objective of my study is to close this gap between theoretical and empirical research. Therefore, I analyze the influence of NSP on individual and collective decision-making empirically. Firstly, I investigate whether the existing theoretical considerations on NSP and their impact are correct. Secondly, I look for additional, yet unknown, patterns to supplement further arguments to the discussion on nonseparability. This complements traditional research as "almost all existing work in analytical political theory takes a behavioral perspective, seeking to explain how people act [and] why societies take the political decisions they do" (Hinich and Munger, 1997, p. i). It allows me to draw conclusions about the impact of NSP on actors' behavior and, therefore, on how decision-making processes in politics can be better understood.

[12]The concept of salience in theoretical models denotes the importance an actor attaches to the issue in question (Hinich and Munger, 1997).

[13]Looking beyond political science, Russ (2011, p. 211) summarized contemporary contributions in the disciplines of biology, psychology, finance and economics with the mnemonic "complexity science has come into being."

1.2 Previous research

The basic principle of nonseparability states that if a connection exists between individual parts of an overarching theme, those parts should not be separated. Far from being an innovative concept, the idea of nonseparability has existed for quite some time. It has also been discussed for its relevance in analytical political research (e.g., Strom, 1990). In this section I review the previous contributions to and applications of the nonseparability principle.[14] My interest was not to focus only on theoretical work, but empirical contributions have been rare. Rather than just referring to existing work, I will also point out the missing elements of NSP and include possible improvements. I argue that investigating NSP can enhance scientific research in many subjects: from referendum and survey research, to organizational theory, and institutional analysis. By mentioning the potential fields of application so prominent at this point, I hope to contribute to the benefit of future research in analytical politics. The consequences of neglecting nonseparability differ from field to field, but in general NSP may distort results and lead to wrong conclusions.

The field which has explored most explicitly the issue of nonseparability is that of survey research. This holds for its measurement as well as for its explicit designation. Therefore, I start my literature review with contributions in this field. The bulk of literature on nonseparability, however, has focused on its potential for strategic manipulation (e.g., Holt and Anderson, 1999). This subject comprises innate political fields such as administrative decisions, international relations, elections and referendums. After discussing this work, I look at the existing (empirical) research on the interplay of procedural rules and NSP. This covers the fields of organizational politics as well as institutional analysis. The inherent logic of nonseparability is clearly linked to central topics of those subjects, e.g., delegation or decentralization. Both concepts are concerned with the separation or unification of decision-making competences. But nonseparability is not only relevant in institutions, it exists within preferences. Therefore, I complete the review by discussing (mainly theoretical) contributions on the modeling of utility functions of individual as well as collective actors.

SURVEY RESEARCH

Surveys rely heavily on the measurement of individual preferences, and it is essential to measure them correctly and precisely. The problem of conditional preferences is discussed in this field in the context of question order effects (Schuman and Presser, 1996).[15] These effects show that the answers to identical questions vary systematically because of the way in which they are ordered. Most contributions look into the effect of rearranging questions. For example, Malecki and Gabel (2007) investigated a change in the

[14]Chap. 2 elaborately discusses the concept of nonseparability including a multitude of examples. In particular, Sec. 2.4 explores the theoretical foundations.

[15]Question order effects are separated into "part-whole" and "whole-whole" effects. The first refers to questions with super- and sup-ordinate parts (e.g., "Should taxes be increased?" - "On what should the surplus be spent") and the second to questions on the same level (Willits and Ke, 1995).

Eurobarometer survey[16] and Siminski (2008) pointed out that question order effects are not necessarily limited to single questions. Instead, "question order has the potential to cause systematic positive or negative bias on responses to all questions in a battery" (Siminski, 2008, p. 477). The literature explains this by default with framing, consistency and salience effects (Schuman and Presser, 1996, p. 12ff).[17]

Lacy (2001a,b) was the first to discuss NSP in relation to surveys. The measurement of separable preferences can be limited to the actors' first preference. Measuring nonseparability requires an evaluation of preferences according to several specifications, making this a disproportionately more complex task (cf. Sec. 2.4.3). Nevertheless, Lacy (2001a,b) investigated nonseparability in respondents' preferences in the United States (U.S.) when answering survey questions. He demonstrated that respondents' preferences with regard to income taxation depended on crime prevention policies; preferences on environmental pollution depended on environmental regulation; preferences on defense spending depended on social spending; preferences on immigration policy depended on the constitutional status of English being the only official language in the U.S.; and so on. Thus, NSP exist in real world political issues[18], and they might affect how people vote, form judgments or make decisions. The studies clearly showed that common survey questions about first preferences alone are insufficient when dealing with NSP. Disregarding conditional preferences at the individual level increases the risk of mistaking a respondent's conditional response for the genuine preference. This leads to misconceptions at the aggregate level of analysis.

Bonica (2012) focused on fiscal preferences using an interactive questionnaire. Respondents were asked to adjust single issues spending levels within an overall fixed budget. the results showed that the preferences on security spending correlate with self-reported ideology. This is in line with the findings of Lacy (2001a) concerning the interplay of preferences on defense and social spending.

As the contributions discussed so far agree that NSP exist, they also tried to measure them. Survey questions are suitable for this task, but NSP are far from being an accepted standard in this field. Too little is known about their impact or importance. It is also necessary to look into research in other areas and to bring the respective findings together.

STRATEGIC MANIPULATION

Most of the theoretical literature on nonseparability has focused on the potential for strategic manipulation when separating or combining the decisions over issues which influence each other (e.g., Ordeshook, 1986). Already Schwartz (1977) set up a general

[16] The Eurobarometer is a regularly and biannual survey conducted on behalf of the European Commission. It reports public opinions across the EU member states and is accessible at `http://ec.europa.eu/public_opinion/index_en.htm`.

[17] Cf. Druckman (2004) for an assessment of political conditions (e.g., elite competition, deliberation) under which framing effects occur.

[18] The contribution of Lacy (2001b) clarifies also that dependencies are very well possible across policy fields.

model to investigate the separation of a bundle of positions into single issues. The model focused on compromises across issues in which "vote trading (log rolling) unseparates separate issues [by] combining positions on distinct issues to form single legislative packages" (Schwartz, 1977, p. 999). He concluded that vote trading can have advantageous as well as disadvantageous influence in the attempt to achieve a Pareto-efficient outcome.[19]

In the case of NSP, and in contrast to a trade, the preference of an actor in one issue changes with the result of a second issue. Probably the most obvious situation in which nonseparability occurs is a dependency of spending preferences. Hinich and Munger (1997) discussed the consequences for two or more public policies when they are subject to the same budget constraint. Given this constraint, the amount of money allocated to one policy program reduces the overall funds still available. Thus, the allocation to one program influences the spending preferences over all remaining policy programs. Whether they increase or decrease depends on whether the programs are substitutes or complements.

Linking single components is by far not limited to budget negotiations. Morgan (1984, 1994) applied bargaining theory to international crises and argued that issue linkage plays an important role in the bargaining process. He identified the incorrect specification of such interconnections as one possible reason for the theory's explanatory shortcomings. In his studies the focus was on three aspects, which so far have not been included sufficiently: the relative salience of issues, the relative power of actors, "and the possibility that the linkage issue raises other issues with which it is inversely related" (Morgan, 1990, p. 329). As this conditionality was not accounted for in the original framework, he judged its setup as "underdeveloped" (Morgan, 1990, p. 312).

In the context of economic sanctions and military coercion, Lacy and Niou (1998a) supplemented the theory of issue linkage. The authors "develop a game-theoretic model of disputes on multiple issues where disputants may have either separable or nonseparable preferences" (Lacy and Niou, 1998a, p. 2).[20]

These various theoretical models of strategic manipulation prove that researchers are aware of the possible influence of NSP. However, their consideration is far from state-of-the-art in any field. At least they provide a solid basis for future empirical tests. I hope that my own study encourages this.

ELECTIONS AND REFERENDUMS

The separation, combination or concatenation of decisions is an important aspect of all democratic decision-making procedures, e.g., when elections and referendums ask for

[19] An outcome is defined as Pareto-efficient or Pareto-optimal if no other alternative is Pareto-preferred to it. That is, no other alternative can make at least one individual better off without making another individual worse off (cf. Dougherty and Edward, 2012, p. 656).

[20] Interestingly, Lacy and Niou (1998a) did not even regard the inclusion of NSP as their crucial contribution. Instead, they argued that the "key feature of the model is that states have incomplete information about each other's preferences" (Lacy and Niou, 1998a, p. 2).

the electorate's opinion. It is inevitable that they enable voters to express their preferences in an undistorted, clear and complete manner.

Benoit and Kornhauser (1991) took a look at assembly elections in which the electorate had to vote for single candidates. The authors pointed out that, in fact, citizens' preferences concerning the constitution of the assemblies as a whole, and not the platforms of the candidates themselves, are essential. To justify the assumption of separable preferences it is therefore necessary to investigate assembly-based and not candidate-based preferences. For this purpose the authors introduced the term "simply voting", which extends the idea of sincere voting, with the assembly level in mind. Yet, they acknowledged that even then "seat-by-seat procedures are efficient or neutral only under extreme conditions" (Benoit and Kornhauser, 2007, p. 1).

Somewhat more sophisticated are election models for double-member districts. Here, Lacy and Niou (1998b) investigated electoral equilibria under the assumption of NSP and simultaneous voting. The inclusion of NSP creates multiple equilibria and favors candidates with extreme positions. These findings were quite different from the then-standard results and models (e.g., Cox, 1984). The authors emphasized that many researchers recognize the pivotal role of the separability assumption, but that until their contribution it was "imposed on preferences in nearly all of the formal models of elections and legislatures" (Lacy and Niou, 1998b, p. 91).

In general, polls are well suited to clarify the impact of interrelated issues. Thus, a simple and single binary vote (whether in an election, referendum or opinion poll) may not be enough if the topic in question contains interrelated issues. This would risk mistaking a conditional response for a genuine first preference (Lacy, 2001a). Conclusions based on such results are biased and inaccurate.

But do such relations really exist? Carrubba and Singh (2004) argued that individuals can have quite sophisticated policy preferences incorporating multiple layers. Using the example of a European Union common defense force, the authors emphasized that ignoring "correlated preferences" leads to falsely-specified empirical models (cf. Carrubba and Singh, 2004, propositions 2 and 3, p. 221). In addition, these secondary effects may also explain volatility in public opinion polls.[21]

Often, a referendum asks its participants for assessments on several issues simultaneously. Brams et al. (1997) argued that such a voting procedure prevents the complete and sincere representation of voters' preference ordering and, ultimately, can cause representation bias. Their work highlighted the importance of procedure in dealing with NSP and suggested alternative aggregation and voting rules. The authors discussed approval aggregation, split aggregation, Borda Count and approval voting. All these rules offer

[21] Cf. Zaller and Feldman (1992) for a more comprehensive discussion whether surveys actually reveal preferences and on "methodological artifacts" as question order effects, semantic priming, consistency effects, framing effects, etc.

the participants more options for expressing their preferences:[22] approval aggregation counts abstention at the same time as a yes and a no vote; split aggregation divides every abstention evenly across all alternatives; Borda Count asks every voter for all desired combinations of issues and awards the number of votes according to the listed order; approval voting enables citizens to vote for all acceptable combinations.[23]

Lacy and Niou (1994, 2000) had similar research objectives when looking into referendums on multiple binary issues. They were concerned that majority rule may select outcomes inferior to the Condorcet winning set when voters have NSP across the issues. The choice may have even been a Condorcet loser or Pareto dominated by other sets.[24] Like Brams et al. (1997) they proposed alternative voting procedures like set-wise voting, issue-by-issue voting and vote-trading. These mechanisms prevent the selection of inferior outcomes and are facilitated by legislatures.[25] This leaves a usual referendums with the ambiguous characterization that it maximizes "the quantity of participants in democratic decision-making but minimizes the quality of participation" (Lacy and Niou, 2000, p. 1).

These insights into referendums were contested by allegations that the undesirable outcomes involve contrived voting situations that would be unlikely to occur in actual elections. Using computer simulations, Hodge and Schwallier (2006) responded to this accusation by investigating the desirability of referendum outcomes. This robustness check also validated the negative effects of unaccounted NSP in randomly generated elections. In addition, the simulations provided support for the claim that certain alternative voting methods can more accurately reflect the will of the electorate.

ORGANIZATIONAL POLITICS

Previously, I discussed the strategic separating of issues; now, I turn to analyzing decision-making structures. Both fields are strongly connected as they ask the same question: what influence is dedicated to the process of decision-making? When organizational units are confronted with decisions on jurisdiction or subsidiarity they are dealt with by delegation, decentralization or specialization (e.g., Shy, 1996). Each organizational form implies the separation of decision-making with respect to either a spatial, temporal, hierarchical or administrative dimension (Bendor and Meirowitz, 2004).

[22] None of the rules was designed to deal with the problem of NSP. Rather, they intend to provide a better opinion assessment of decision-making with multiple alternatives. Yet, nonseparability links otherwise separate issues thereby generating a plurality of alternatives. For example, in the case of the aforementioned exquisite dinner the combination of the binary decision for each wine (white or red) and main course (fish or venison) results in four possible variants of which one is the result. Thus, the voting schemes can be helpful in the identification of NSP.

[23] Hodge (2006) offered a more elaborate overview on alternative voting rules and discusses (non)separable preference orders in the broader context of combinatorial questions.

[24] Such outcomes are called "multiple-election paradoxes" (Xia et al., 2011).

[25] Lang and Xia (2009) agreed that to allow voters to express their full preferences on the set of all combinations of values may prevent multiple-election paradoxes. But this is practically impossible as soon as the number of issues or the size of the domains increases above a few units. By relaxing separability restrictions, they identified a practicable compromise. The authors therefore applied a sequential composition of local voting rules related to the setting of conditional preference networks (CP-nets).

Much of the literature in the field of organizational theory investigates these organizational tools. Their benefits, implications and requirements are discussed and known in detail (cf. also Sec. 2.3). The impact of each decision rule is assessed with respect to the output performance of the respective institution. This search for the optimal organizational structure is important, e.g., for public administration, the industrial sector and research institutions (e.g., Olsen, 2007).[26] If the separation involves decisions over multi-dimensional issues characterized by conditional dependencies, this leads to unwelcome side effects and a loss of efficiency.

A special topic is the growing field of organizational politics, which at its very core is interested in the effects of political tactics (coalition building, sanctions, ingratiation, etc.) on organizational functioning (for a comprehensive overview cf. Ferris and Hochwarter, 2011).[27] Its two most common approaches are the investigation of the effect of political behavior on organizational as well as individual attributes (e.g., work processing and career success) and to comprehend the reactions of individuals to perceptions of organizational politics (cf. Witt et al., 2000, p. 42). While the first aims for a quantitative (objective) research implementation the second views responses as based on an individual's (subjective) perception of reality (Lewin, 1936).[28] As it holds true for organizational theory in general, the assessment of behavior must be based on an accurate specification of dependencies. Otherwise, the conclusions reached are distorted.

INSTITUTIONAL ANALYSIS

Institutional analysis is closely related to organizational theory. Many contributions have long investigated the trade-offs inherent in every institutional reform debate. Tensions between stability, efficiency and transparency are at the very heart of every negotiation. Prominent examples are the federal reform of Germany (e.g., Burkhart, 2009) as well as the reform of European treaties (e.g., Finke, 2009). Consequences of reallocating agenda control or decisions powers must be evaluated correctly, otherwise, an assessment of reforms achieved is not possible.[29] Therefore, interdependencies between different levels of government or political fields must be captured accurately.

[26] For the field of politics Moe (1990, p. 39) pointed out that it is "the structural means by which political winners pursue their own interest".

[27] Vredenburgh and Maurer (1984) provided an introduction into the field of organizational politics by reviewing previous concepts. Incorporating them into a coherent process framework for understanding individual and group political activity the authors synthesized the following definition: "Organizational politics (a) is undertaken by individuals or interest groups to influence directly or indirectly target individuals, role, or groups towards the actor's personal goals, generally in opposition to other goals, (b) consists of goals or means either not positively sanctioned by an organization's formal design or positively sanctioned by unofficial norms, and (c) is objective and subjective in nature, involving real organizational events as well as perceptual attributions" (Vredenburgh and Maurer, 1984, p. 50).

[28] Witt et al. (2000) scrutinized the impact of participation in decision-making: does it matter for job satisfaction i) if decisions are reached in consensus with the supervisor or ii) if orders are given without contestation. Here, satisfaction depends "on a judgment by the individuals as to whether a perceived behavior is within parameters of sanctioned behavior" (Witt et al., 2000, p. 343) or not. Such judgments are highly subjective.

[29] Shepsle and Weingast (1984, p. 49) demonstrated "that institutional arrangements, specifically mechanisms of agenda construction, impose constraints on majority outcomes."

Interestingly, there are fields of research closely considering dependencies within their field as well as with adjacent areas. Those are not necessarily denoted with the term "nonseparability" but research considers the relationships explicitly. The current discussion on the future of the German pension system may serve as a first example of a field with strong interdependencies, and of a policy area characterized by various reform initiatives. In this area, the campaign for the parliamentary elections in Germany in September 2013 brought about a new wave of reform proposals. While they comprise a variety of regulatory and legal aspects, Börsch-Supan and Gasche (2013) conclude that all propositions are nevertheless just "moving injustice from one social security system into another" and illustrate this by a tax-financed basic pension which burdens labor costs and reduces employment. The strong ties between the individual components of the German welfare state make changing a single part very difficult.[30] This points out that (many) political decisions influence adjacent areas.

If one takes a closer look, the federal system of Germany also serves well as a more in-depth case.[31] The intertwining, i.e., the interdependence, of national and federal powers in Germany is well-known and well-researched (Wachendorfer-Schmidt, 2005, p. 11). Its disentanglement was expected to increase the efficiency as well as the transparency of politics (Zohlnhöfer, 2008, 2009). The reform aimed at key areas of the political order, such as legislative power distribution, division of labor within the administration, budget responsibilities, etc. The reallocation of agenda control was another key part, and related power struggles were the main reason for the delays and difficulties of the reform (Benz, 2005).

Interdependencies in political systems are not per se good or bad (cf. Wachendorfer-Schmidt, 2005, Chap. 6, p. 377ff), but they lead to certain behaviors that can be assessed properly only if one knows of these dependencies and takes them into account (e.g., Ganghof and Manow, 2005). This also holds true for reforms, e.g., when the possible consequences of reallocating legal responsibilities are evaluated. A correct assessment is not possible without considering the structure of the system. Therefore, interdependencies between different levels of government or political fields must be captured to evaluate the mechanisms of politics.

Empirical contributions in the field of institutional analysis concerned with nonseparability focused on changes in the policy outcome when conditional decisions are split up by institutional requirement (e.g., separation of power or regional jurisdiction). These institutional rules range from agenda-control over decision competencies to the general institutional environment. To that point, nonseparability is just one of many factors influencing the final outcome as it relates to the overall structure (McCubbins et al., 1989) or the so called "structure-induced equilibria" (Shepsle and Weingast, 2004). Yet, insti-

[30]For a comprehensive overview on the development and current status of the German welfare state cf. Schmidt (2012).

[31]For an up to date overview on general theories about multilevel systems cf. van Deemen (2009) and for a more international perspective cf. Kaiser et al. (2013).

tutional rules are far from unbiased; they provoke strategic behavior hampering equilibrium solutions (e.g., van Deemen, 2009).

A well-known characteristic of U.S. politics is the principle of divided government (Fiorina, 1992b). Smith et al. (1999) analyzed voting behavior in the U.S. and focused on the question of whether the voting decision for the president is dependent on the current legislative majority. Using national and statewide surveys, Craig et al. (2006) had the same objective. While no study found a decisive answer, the strength of the effect depended on voters' level of information (Fiorina, 1992a), i.e., if they knew the current and probable future majority in the House of Representatives and the Senate.

Strategic voting is an essential part of all modern electoral systems, and the respective tactical moves may include very sophisticated considerations. Shikano (2009, p. 271) proposed for Germany's national "mixed-member electoral system" an interaction effect between voting decisions and party competition.[32] He argued that expectations concerning the national level proportional representation influence the electorate to vote strategically in the plurality tier. Since the relevant variables were not available through survey data, he used computer simulations. These simulations were most appropriate in explaining short-term linkage between the two tiers.

Thiem (2009) investigated agenda-setting and roll-call vote (RCV) requests in the European Parliament (EP). The EP offers a unique case, as a member of the European Parliament (MEP) is considered to be an agent to their national party and their European group simultaneously. Both bodies are able to discipline their MEPs. The author argues that there are theoretical reasons to question the connection between strong European group leadership and unity among the factions in roll-call votes. Because disciplining abilities of the European groups are limited, party unity is more likely a condition for and not the result of a RCV. Thus, the role of national parties may, so far, have been underestimated. This also implies that RCV in the EP serves mostly as a signaling device (towards national constituencies) and is not constituted for disciplining purposes.

These contributions show that the influence of the institutional structure or single organizational peculiarities can be misinterpreted if conditionalities are not assessed correctly. Behavior due to a specific dependency may be attributed to an aspect which is not at all involved or responsible.

ANALYTICAL POLITICS

Conditionality is not only relevant in combination with one institutional arrangement or specific agenda-setting[33]. Analytical politics covers a wide field of topics in which it aims to analyze and explain political behavior (cf. Hinich, 2008). It focuses on the identification of empirical regularities.

[32] Cf. Soberg and Wattenberg (2001) for a comprehensive overview on mixed-member electoral systems and their characteristics.

[33] Wilson (2007) provided an extensive overview on the topic of agenda-setting in political science and economics.

The concept of abstraction is a key aspect of all analytical research as it transforms complex (policy) processes into manageable models. Those models serve as "stylizations meant to approximate in very crude fashion some real situation" (Shepsle and Bonchek, 1997, p. 9) and incorporate preferences of the actors involved, in the form of utility functions. NSP alter individuals' utility function, exhibiting effects in all configurations. This must be considered at an early stage when setting up analytical models containing the specified institutional arrangement of interest.

A well-known and often used tool in the field of analytical politics is spatial modeling (e.g., Davis et al., 1970; Krehbiel, 1988). The operationalization of spatial models consists of formulating indifference curves (for a more detailed discussion cf. Sec. 2.4.2). These curves rely on correctly specified utility functions, i.e., the correct assessment of individuals' preferences.[34] If the curves are inaccurate, the results obtained from the spatial analysis are implausible. Different models can neither be compared, nor crucial aspects of (political) behavior identified.

In general, models in the realm of analytical politics describe individual actors and their behavior. But this does not mean that each actor must actually be an individual person in the real world. Theoretical models allow incorporating administrative bodies, companies, legal persons or societies as well. The models abstract from the type of actor and operationalize both individual and collective protagonists as unitary actors (e.g., Baron and Ferejohn, 1989). Here, Le Breton and Sen (1999) made a first step by discussing social choice functions[35] (SCF) which were "separable only with respect to the elements of some partition of the set of components and these partitions vary across individuals" (Le Breton and Sen, 1999, p. 605). Thus, the simplifying separability assumption was abandoned, at least in part.

Over thirty years ago, Stiglitz (1972) examined the importance of legislation on bankruptcies, take-overs, and the financial policies of companies. He showed that the real-term decisions of a firm (i.e., the manufacturing sector) are not separable from its financial decisions which depend on its debt-equity ratio. Previous studies had claimed that, if there is no chance of bankruptcy, financial policy has no effect on the value of the firm, implying that there is no optimal debt-equity ratio. Thus, financial regulation may not only affect a company's legal status but also its every-day decisions. During the current global financial crisis, this insight is more important than ever. It highlights the importance of regulators in financial markets and their scope of action (cf. Doina and Nicolae, 2011).

Finke (2009) argued that governmental preferences with regard to expanding the EU's jurisdiction depend on the decision rules applicable inside the EU's decision-making bod-

[34] While recent contributions increasingly paid attention to preferential distributions and actor-specific characteristics (e.g., salience, Benoit et al., 2005; Slapin and Procksch, 2010) they stick with separable preferences as default.

[35] Le Breton and Sen (1999, p. 605) defined a SCF as "a mapping which associates a social alternative with every profile of individual preferences defined over a set of social alternatives. The value of a SCF at any profile is to be viewed as the 'optimal' or most 'desirable' outcome in that state of the world."

ies and vice versa. Whereas some governments preferred further integration only if they were to maintain their veto power (e.g., unanimity rules in the Council of Ministers), others preferred further integration only if decision-making procedures were to become more efficient (e.g., lower voting thresholds). The underlying logic can be summarized as public goods inhibiting economies of scale, set against individual policy losses due to heterogeneity. The author tested his hypothesis by applying a two-stage item response model to data on governmental reform positions.[36]

Analytical models rely on the implementation of individual utility functions, much in the same way as surveys rely on the measurement of individual preferences. Its underlying pattern matters for the assessment of behavior. In both cases, only the correct operationalization establishes the nonnegotiable and fundamental condition for proper scientific research. Thus, NSP represent not a specific obstacle of a single field of research. This is also evident in the wide-ranging fields covered in this summary.

1.3 Focus of the present study

In short, this study faces a trade-off. Every contribution to analytical research must strike a balance between scientific interest and existing resources, between including more details versus keeping modeling within manageable limits. I focus on the negligence of NSP. More precisely, I ask whether analytical research should override them and keep the analysis simple but potentially biased, or include nonseparability together with its complex modeling requirements.

Whereas the theoretical concept of NSP is relatively old, empirical tests thereof are rare. Except for the literature discussed in Sec. 1.2, almost all research projects use the simplifying assumption of separable preferences as their default. This may lead to distorted results and invalid conclusions. Of course, many studies which exclude NSP may be right in doing so. Yet, the nearly unambiguous exclusion from analytical political research can hardly be justified just because it is traditional to do so.[37] In other words, I do not argue that every research project should engage in measuring or operationalizing nonseparability. This depends in each case on the specific research question and design. But scientific research standards can only be maintained when the role and relevance of NSP have been duly addressed.

Analytical politics dismantles challenging complex political phenomena into feasible components, using models as "internally consistent bodies of theory that describe human behavior" (Hinich and Munger, 1997, p. 3). Applying these models enables the analysis of

[36] Shu (2003) investigated a related field when looking into NSP in public opinion polls and referendums on European integration.

[37] Benoit and Laver (2007c, p. 31) made a similar statement to another quasi-standard of political science research, the usage of Euclidean utility functions when modeling political decision-making. I will come back to that discussion in Sec. 2.4.3.

intricate questions by reducing complexity through abstraction from reality. If one utilizes this research paradigm and accepts the general principle, the degree of abstraction still remains controversial. What details can be ignored, which features omitted, which simplifying assumptions can be made without losing validity? In general, this depends on whether or not the aspect in question is of significance for the findings of analytical political science research. My analysis reveals the necessary requirements and knowledge gained with respect to NSP. This enables me to make the significance assessment based on a sound foundation.

It is far beyond the scope of a single study to look into every individual sub-field of analytical politics. I restrict myself to a single but central aspect, the formulation of theoretical models, which form the very core of analytical politics. Here, their purpose is to explain collective decisions. Yet, all collective action rests upon individual behavior. Therefore, the models implement human behavior in a computable way; it goes without saying that it is essential to model the underlying behavioral concepts properly. Only then can the results of the empirical analysis be obtained in an intelligible and reliable manner. Ganghof and Manow (2005, p. 11) pointed out that in addition to the disclosure of behavioral assumptions "the focus [of rational choice] rests on analytically distinguishable mechanisms and systematic institutional effects." To uncover such patterns it is important to formulate the underlying function, i.e., the formal model, punctiliously. But the unambiguous assumption of separable preferences, usually without taking other possibilities into account, is far from appropriate.

The relevance of my argument to consider NSP can be summarized by two key questions. Firstly, whether and, if so, how much the inference of analytical research suffers when falsely assuming separable preferences? For example, are policy recommendations about which decision process should be used (e.g., open vs. closed rule) still valid? Secondly, whether and, if so, how nonseparability affects individual and collective behavior? Preferences determine the action of every individual. Yet, behavior is also shaped by institutions (Plott, 1979); perhaps more in the domain of politics than anywhere else.[38] Thus, what influence does nonseparability exert under different institutional arrangements?[39] To answer these questions, I will use both field and laboratory data, and examine collective as well as individual behavior.

Initially, I intend to establish the relevance of NSP. For this, a review of previous research will answer the question of whether the extra expense of considering nonseparability is worthwhile. I have chosen the field of legislative decision-making, where conflicts over law proposals are frequently conceptualized as being multidimensional (e.g., Thomson

[38] The separation of powers is one of the most fundamental concepts of political theory (de Groot-van Leeuwen and Rombouts, 2010).

[39] Institutions affect behavior in a variety of ways (cf. Shepsle and Bonchek, 1997). For example, democracy encourages cooperation (Dal Bo et al., 2010), procedural justice determines the evaluation of a given outcome (Lind, 1988) and information aggregation is strongly related to the applied voting rule (Morton and Williams, 1999). My investigation will therefore also look into the relationship of nonseparability with such further aspects.

et al., 2006), but studies in this field rarely engage in a discussion of the potential interrelation between issues with respect to actors' preferences. I therefore begin my empirical investigation by examining the impact of NSP on the performance of several models of legislative decision-making. Using empirical data, I endeavor to determine the relevance of neglected nonseparability. The computations also demonstrate the problematic data collection at the individual level in the field.

In a second step, I bring the discussion a more comprehensive level. I am interested in understanding the implications of NSP on individual and collective decision-making in general. The focus is on universal patterns. Thus, I refrain from using observational data from only one specific environment. Instead, I turn to the general and universal framework of laboratory experiments.[40] My experiment compares individual and collective votes regarding a problem, characterized by NSP. The records offer insights into the dynamics of the decision-making process (cf. Kocher and Sutter, 2005). The decisions are reached in the group at large or in situations in which the group is split. Comparing those, I review existing hypotheses and look for still unknown patterns.

This approach underlines very well the interdisciplinary character of this study. On the one hand, the use of laboratory experiments and rational choice are far more common in economics (cf. Sec. 4.1). On the other hand, the juxtaposition of procedures is representative of political science (e.g., Steunenberg, 1994; Gailmard, 2009). Reviewing the literature of the two fields, Mertins (2008, p. 27ff) concluded that there had been rather little research on the dynamics of procedures in economics. That may sound harsh, but it just depends on the field to which field one attributes such excellent contributions as, e.g., the work of Elinor Ostrom (most prominent Ostrom, 1990).[41] Nevertheless, at least compared to political science the research in "economics has devoted little attention to whether the type of decision maker [group or individual] matters for economic decisions" (Kocher and Sutter, 2005, p. 200).

Finally, I bring the different findings together. The consequences for analytical research dealing with NSP depend on two aspects. Firstly, the effort necessary to incorporate NSP must be assessed. Secondly, researchers must have an understanding of how much knowledge is to be gained by expanding their analysis. The final part of this study evaluates these aspects and weighs them against each other.

1.3.1 Research design

This study applies rational choice theory. After all, this theory and its assumptions are the reason for my research in the first place. Models of rational choice commonly employ utility functions based on preferences to describe human behavior. To specify such

[40] This is straight in line with Plott (1991, p. 902) who argued that "laboratory methodology involves a shift from a focus on particular economies as they are found in the wild to a focus on general theories, models and principles that govern the behavior of economies".

[41] For an overview of her social science masterpiece cf. Nutzinger (2010).

functions requires the application of simplifying assumptions, and a delimitation of the object of study. Thus, the question - as well as the investigation - of the nonseparability assumption is inherent to rational choice. Most importantly, current contributions emphasize institutional factors and cognitive aspects of behavior explicitly (e.g., Lane and Ersson, 2000; Simon, 1982).

Rational choice theory aims to explain social behavior, while its objects of study are individuals (Shepsle and Bonchek, 1997, Chap. 2). This is based on the insight that all social relations must be derived ultimately from individual decisions (Braun, 2013, p. 164). This methodological individualism has its limitations, as the aggregated collective preferences do not necessarily correspond to individual preferences (Coleman, 1990). However, rational choice focuses on the preferences and beliefs of individuals, and assumes that they behave rationally and in their self-interest. This means they choose between available alternatives by ranking them according to their preferences and beliefs. For this purpose, the approach utilizes "the formal precision instruments of micro-economics, decision theory and game theory" (Holzinger, 2009, p. 544). In order to act rationally, each alternative is assigned a relative value. Applying a "first generation" version (Braun, 2013, p. 164), it is possible to make the right choice, even in very complex situations.[42]

I agree with Morton (1999, p. 24) that "empirical analysis has, can, and should be used to empirically evaluate formal models in political science." In this field, rational choice theory is a methodically disciplining access. Empirical research that uses this approach is compelled to disclose its basic assumptions about actor behavior. This enables a thorough analysis of the logical framework and prevents logical fallacies (cf. Bueno de Mesquita, 2004). Schotter (2006, p. 500) argued that "the combination of rational choice models and empirical support is the only way to gather knowledge in the social sciences". But rational choice theory has often been criticized.[43] Rather than providing an in-depth literature review, I follow up with the main criticisms of the theory: to much attention is given to the formation of preferences, while institutional contexts are largely ignored. Furthermore, a series of implausible assumptions and unrealistic constraints shows limited understanding of the human individual (cf. Green and Shapiro, 1994; Taylor, 2006). Nevertheless, it has achieved wide acceptance in economics in recent history (Becker, 1976; Franz, 2004), even if it is still very controversial in sociology and political science (cf. Kirchgässner, 2008; Opp, 2004; Vanberg, 2002).[44]

One particular point of criticism is the traditional assumption that a person can at any rate order their preferences amongst a variety of alternatives, and chose the best option. The

[42] Simon (1993) referred to this degree of complete rationality as "Göttlichkeitsmodell" (i.e., model of divinity). Here, four criteria of rationality must be fulfilled (von Neumann and Morgenstern, 1944): completeness, transitivity, continuity, and independence.

[43] Gilboa (2010) offered insights to the rational choice paradigm in general. He also provided an overview of its development due to criticism of the economic nature of the approach.

[44] Holzinger (2009) concluded that in addition to the difference between disciplines a difference between the German and international approach exists. She demonstrated this by going through the single sub-fields of political science such as political theory, international relations, etc.

presumption that a person can keep track of all the possibilities to choose from is quite strong (Sen, 1997).[45] The first fundamental contributions (e.g., Arrow, 1950; Niskanen, 1971; Olson, 1965) are today an integral part of political science literature.[46] But following its strong presumptions, these early models could not explain altruistic or collective behavior within the rational paradigm. Fortunately, since the 1950s and 1960s a growing amount of research has established a much broader framework (e.g., Kahneman et al., 1982; Sen, 1995).[47] While some years ago Frohlich and Oppenheimer (1996, p. 118) argued that "concern for fairness [are] particular problematic for economic theory", nowadays, models can incorporate social concerns (e.g., Ockenfels, 1999). The parsimonious setups were continually expanded, which made rational choice into a much more hard-headed tool when investigating behavior and decision-making (Wandling, 2012).

Consideration of psychological research has distinctly improved the realism of analytical research (e.g., Quattrone and Tversky, 1988). Researchers now are able to incorporate into their theories the idea that "it is sometimes misleading to conceptualize people as attempting to maximize well-defined, coherent, or stable preferences" (Rabin, 1998, p. 1).[48] In this context, Schotter (2006, p. 500) concluded that "rational choice theory does not provide us with the truth about human behavior but rather with a very compelling platform with which we can seek the truth." Undeniably, rational choice theory is neither always the proper nor the only possible theoretical framework for scientific research.[49] But according to Schotter (2006, p. 507), it is the "theory that offers us a platform to understand what is failing in our assumptions and to change those elements without throwing the rest of theory away." Step by step, the theory became more realistic and less idealized by incorporating previously omitted details (Ariely, 2009).

NSP have been disregarded for a long time even though an astonishing amount of work has been done in two strands of analytical research. Recently, political scientists have begun to focus on the issues of causal complexity.[50] These studies have improved the capacity to empirically test for conditional effects (respectively "causation", cf. Brady, 2008; Gerring, 2005), and to disentangle the sufficient and necessary conditions for the

[45] According to Sen (1977), a person who depicts these abilities would have to be completely rational. But, even then an individuals' rational or egoistic action does not per se lead to a general equilibrium. Still, this rationality means that the person would not reveal inconsistencies in their choice behavior. On the other hand, they would not be able to consider the total variety of alternatives carefully, thus, leading to the conclusion that the "purely economic man is indeed close to being a social moron" (Sen, 1977, p. 336).

[46] Braun (2013) discussed these first contributions to rational choice theory in detail.

[47] Mitchell (1999) reviewed in detail the then valid tenets and methods of political science from 1950 to 1970.

[48] In this study I will discuss current approaches which take into account that individuals are not purely selfish (for a literature overview cf. Sobel, 2005). This development is not entirely contrary to the view of classical economists. Following Smith (1759, p. 1) how "selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it."

[49] Concerning alternative methods, King et al. (1994) offered a comprehensive volume on scientific inference in qualitative research.

[50] Pearl (2009) discussed the paradigmatic shift that was necessary to enable the causal analysis of multivariate data. In short, he concluded that the data or distribution of data alone is not enough. Causal questions require some knowledge of the data-generating process (DGP).

occurrence of political events (e.g., Braumöller and Görtz, 2000; Brambor et al., 2006; Ragin, 2006). At the same time, increasing progress has been made to operationalize spatial models of legislative politics (e.g., Laver and Schilperoord, 2007; Martin and Quinn, 2002; Slapin and Proksch, 2008).

The improvements in both fields laid the foundation for enabling the examination of nonseparability in an exact and precise manner. By no means should these contributions now be ignored. Rather, future research should be adjusted to include the consideration of NSP. This follows Ostrom (2006, p. 8) who advocated that "one does not just toss out a theory that has proven valuable in many settings because it does not work well in others. Many efforts to broaden the theory are well underway, and it will continue to be usefully employed to address many interesting questions in competitive situations."

1.3.2 Why a laboratory experiment?

I investigate the effect of nonseparability on individual and collective decision-making using experimental methods. Krehbiel (1986, p. 547) already summarized the obstacles in assessing collective decisions: preferences are difficult to measure, institutions are remarkably complex, and strategies are highly diverse. The measurement of NSP is especially challenging (cf. Sec. 2.4) and the same is true for operationalizing it. But in an experimental environment the preferences of the subjects can be induced corresponding to the research question through monetary payments (McDermott, 2002, p. 326). The participants are remunerated for their participation and their performance in the experiment. The amount of the payment depends on their decisions made during the session. By defining the payoffs the experimenter obtains control over the incentive structure of the participants.[51] It is possible to set up the (nonseparable) decision problem in the way most applicable to the research interest (Webster and Sell, 2007, Chap. 9). In other words, I expose the participants to a decision problem affected by nonseparability and observe their reactions. This saves me from starting an onerous search for such situations.

Experiments enable the analysis of each step of decision-making through monitoring approaches (Ordeshook and Winer, 1980, p. 730). This is often not possible using observational data, which yields only the outcome of a decision-making process (Schnapp et al., 2009). Therefore, the behavior of every individual that takes part in a collective decision-making can be scrutinized. The possible juxtaposition of individual and collective levels allows the disentanglement of additional research questions (Halfpenny and Taylor, 1973).

Going into more detail, I use a computer-based laboratory experiment. Compared to a field experiment, the high degree of environmental control enables me to secure a high

[51] The assumption that preferences can be induced by financial incentives follows the axiom of "nonsatiation". This axiom states that individual utility is a monotonically increasing function of the monetary reward (Smith, 1976; Smith and Walker, 1993). In other words, the participants will act according to the given earning opportunities.

level of internal and construct validity (Morton and Williams, 2010, p. 192ff). This emphasis on validity is important, because NSP may appear to many people as quite an abstract feature. Furthermore, collective decision-making involves interaction between subjects (e.g., Butler and Camerer, 2005, p. 12ff). The laboratory facilitates the control of communication (either allow no communication, or prescribe how to communicate) and the necessary assurance that all subjects are taking part at the same time (Cason and Mui, 2007).[52] Because of these characteristics, laboratory experiments have matured over the last two decades and achieved recognition as an approved method for the validation of theoretical models (Palfrey, 2009).

The laboratory allows me to induce NSP in a controlled environment, but comes with the usual trade-off between internal and external validity (Morton and Williams, 2006, p. 8ff). The main disadvantage of experiments is their low external validity, specifically the poor generalizability of the results.[53] The approach should thus not be used for single case studies, but for "theory development, testing and refinement" (McDermott, 2002, p. 341).[54] This matches the aim of my research, as I scrutinize the effects of wrongly specified utility functions. The main concern of my experiment is to examine existing but untested hypotheses empirically (cf. Sec. 2.5). In addition, the excellent monitoring opportunities enable the search for behavioral patterns so far not discussed in relation with nonseparability. This supplements prominently the consideration of their relevance.

Already Smith (1973, p. 3) pointed out that the "concepts of utility are not only basic to microeconomics theory; they are also basic to experimental methodology in general." Thus, rational choice theory and the experimental method complement each other. Furthermore, I can rely on a rich tradition of experiments (cf. Sec. 4.1.2) as well as other research designed to study collective decision-making in general, and committee decisions in particular (e.g., Barbera et al., 2005; Huitt, 1954; Keiding and Peleg, 2001). My work is part of a body of contemporary political science research to which the contribution of experimental methods is still growing (cf. Morton and Williams, 2010).

1.4 Outline of the study

The study is divided into three main parts. The first part clarifies the concept of NSP and emphasizes its relevance. In Chap. 2 various short examples and simple case studies demonstrate the existence of nonseparability in everyday situations. The chapter also covers extensively the (previous) theoretical modeling of NSP. In addition, I add further

[52] In addition, it is much easier to observe which person actually participates. In online surveys always an element of uncertainty remains about who has completed the questionnaire (cf. Eckel and Wilson, 2006).
[53] Morton and Williams (2010) pointed out that experimental findings need to be confirmed by additional contributions. For example, further experiments with variations in design, subject pool and treatment.
[54] Schram (2005) provided an in depth discussion of the artificiality criticism on laboratory findings. He pointed out that it is necessary to distinguish between various goals of experimental contributions and that, e.g., theory testing is affected by this in a completely different way than theory development.

arguments to the theoretical concept that will be scrutinized empirically in later chapters. Chap. 3 presents concrete evidence for the significance of NSP. Here, I conduct an empirical study on the influence of NSP on (spatial) models of legislative decision-making. As a data base I use the widely known DEU project (Thomson et al., 2006). The chapter exemplifies the shortcomings of traditional research methods for analyzing nonseparability. As empirical research on NSP is so far still sparse, it proves the necessity of further investigations.

In the second part, I investigate the impact of NSP on individual and collective decision-making within the framework of different institutional arrangements. As the measurement of separable preferences itself is hard enough, to include conditional dependencies is beyond the manageable scope of traditional research techniques. Thus, the four chapters of this part focus on a laboratory experiment. Chap. 4 lays out the experimental design in detail. It also contains a review of laboratory research in political science, and places the current experiment into that context. The respective sections elaborately describe the design decisions made as well as their implementation. Chap. 5 presents the results of the experiment on the aggregate level. I investigate the influence nonseparability has on the collective outcome. The results are then further scrutinized in Chap. 6, where I turn to the individual level. I analyze the single votes of every committee member to identify the determinants for the individual choices. Subsequently, Chap. 7 reports the results of a post-experiment survey. This adds insights from the participants' perception of the experiment. The questionnaire surveys the subjects about their decision rules and important characteristics of alternatives they have chosen.

The third part concludes my study and is concerned with the consequences for scientific research when encountering NSP. It links the previous theoretical arguments and the empirical results. The rational choice framework prescribes ambitious requirements when condensing a complex reality into theoretical models. Those are even more demanding when dealing with nonseparability, but still the capabilities of models are limited. I discuss this interrelation in detail in Chap. 8. The chapter merges my various findings and offers guidelines for whether and how NSP should be taken into account. It also places my contribution into the current state of scientific research. Finally, Chap. 9 concludes the study with a summary of the previous chapters and an overview of possible future research.[55]

[55]Overall, a multitude of computations was necessary to achieve this varied research design. Sec. A.1 provides an overview of all software used in this study and offers a short description of their application. All programs are listed including their version numbers and relevant functions.

2 What are nonseparable preferences?

In Chap. 1 I introduce the central idea of NSP. Based on this lead-in, the following sections put nonseparability into a broader context. This is necessary to substantiate its significance. Many readers might now expect a highly technical discussion of a generally rare phenomenon; however, NSP are quite the opposite. They are not at all rare. That is the first point I make in this chapter.

In the following I illustrate plainly the underlying idea of NSP. For this purpose culinary examples, as in Sec. 2.1, are widely popular (Lacy, 2003). Moving from hypothetical to empirical examples I depict short case studies verifying the nonseparable character of actual political problems in Sec. 2.2. Those include various fields of political science, such as public opinion surveys or legislation procedures. The illustrative examples were chosen because of their high public salience (in Germany) and the different administrative levels of government (European, federal, state and local).[56] All these instances show that NSP are highly relevant in the context of organizational structures, i.e., decisions about delegation, decentralization or specialization. I therefore summarize the implications for these administrative instruments separately in Sec. 2.3. Next, Sec. 2.4 provides an overview of the various theoretical contributions to the concept of NSP and looks thoroughly into its theoretical foundations. Finally, Sec. 2.5 brings the different arguments together, and points out the consequences of neglecting NSP. Those are tested empirically in the following chapters.

2.1 The general principle

A short and simple definition of NSP reads as follows: An actor's preferences over two issues are considered nonseparable if the outcome on one issue alters their preference ordering over another issue and vice versa (Enelow and Hinich, 1984, p. 18ff). Also, in case each issue is decided separately, an actors' utility depends on the combination of both decisions. While this shortcut is rather brief, there is a lot more to be said about NSP. They are not far away from decisions we make every day.[57]

[56] Admittedly, one reason was also their local proximity to my affiliation at the University of Heidelberg.

[57] Dixit et al. (2009) argued that people constantly interact. In their view they "play games of strategy all the time" (Dixit et al., 2009, p. 3). Here, the term strategy refers to situations which "are distinguished from individual decision-making situations by the presence of significant interactions among players" (Dixit et al., 2009, p. 41). Such real life interactions may well be affected by NSP.

Beginning with such an everyday decision, imagine three couples spending their holidays together in a rented apartment. One evening, the couples decide to have dinner together. How will they organize their meal? One option is that all six go shopping together. This, however, would entail a time consuming exercise in deliberation inside the mall. Hence, they will resort to delegation, which leaves several options. For example, one of them could be appointed chef. Depending on her or his skills, this may or may not be an excellent deal for the remaining five. Another option is that one couple buys the food for the main course, the second is responsible for the drinks, and the third couple is responsible for the dessert. Or maybe the men could get the food and the women the drinks. The problem is that the participants' utility depends on the combination of drinks and food. Imagine that all shopping activity takes place simultaneously (to maximize the time spent together at the beach). Assuming that communication is impossible (the thick walls of the mall are blocking the mobile network), both shopping delegations face a high level of uncertainty with respect to the others' choice. In a worst case they may end up with sweet white wine to venison. How does each of the delegates handle this uncertainty? Given the structure of the "game" (both delegations make their decisions by simple majority), even the knowledge of the others' preference does not solve the underlying uncertainty problem, i.e., there is no equilibrium in pure strategies.

The daily commute to work is another common case to exemplify NSP.[58] Imagine that as a commuter you have two options, going by car or going by bike. In general you prefer going by car, as it is more convenient and faster. This is your everyday first preference. Yet, due to an oil crisis, gasoline gets rationed. The gas station in the neighborhood is already completely out of gas. Now, of course, the bicycle is the preferred means of transport. Otherwise you would end up sitting in a "dry" car. In this example, gasoline and car are complements. Therefore, they ought not to be separated.

Moving the set up closer to a political environment, consider a case in which the Ministry of Transport decides over several alternative routes for a new railway line connecting two cities. At the same time, the environmental administration discusses a reform of its "Biodiversity Act". The exact provision of this act will affect the calculation of expenses for each alternative railway line (and thus the Ministry's preferences). Add to this constellation a construction authority planning to re-regulate the security standards of railway tunnels. Under these circumstances the cost-benefit calculation for the competing routes is highly uncertain. Other examples include the dependency of preferences for delegating competencies to an agency on the distribution of power in the agency's executive board. Another scenario to consider involves preferences to liberalize certain industries that may be conditional upon regulatory safeguards and financial compensations intended to protect the national champion against the adverse effects of competition.

As mentioned before, there are various fields in which one can find dependencies. This prevalence leads to the question how rare NSP are at all? Lacy and Niou (1994, 2000) con-

[58] This example is borrowed from Ted Bergstrom and his lecture on separable preferences (Bergstrom, 2011).

cluded that the separability, and not nonseparability, of preferences may be considered a rather restrictive assumption.[59] Thus, NSP should resemble the standard assumption. Consider the choice over two dichotomous issues. Theoretically, we could encounter 4 different issue combinations resulting in 24 possible preference orderings. Of these, only 8 can be derived by assuming separable preferences, excluding $\frac{2}{3}$ of possible preference orderings. This is shown in Tab. 2.1[60].

Table 2.1: Separable and nonseparable preference orders

EXPLANATORY NOTE
The table lists the possible preference orders for two dichotomous issues. A rule of thumb allows us to determine if an order is separable or not. For two dichotomous issues a preference order is separable if and only if the most desired and the most disliked issue combination form exact opposites as, e.g., YY and NN or YN and NY.

	Separable orders		Nonseparable orders
1.	YY > YN > NY > NN	1.	YY > NN > NY > YN
2.	YY > NY > YN > NN	2.	YY > NN > YN > NY
3.	NY > NN > YY > YN	3.	YY > NY > NN > NY
4.	NY > YY > NN > YN	4.	YY > YN > NN > NY
5.	NN > YN > NY > YY	5.	NY > YN > NN > YY
6.	NN > NY > YN > YY	6.	NY > YN > YY > NN
7.	YN > NN > YY > NY	7.	NY > NN > YN > YY
8.	YN > YY > NN > NY	8.	NY > YY > YN > NN
		9.	NN > YY > NY > NY
		10.	NN > YY > YN > NY
		11.	NN > NY > YY > NY
		12.	NN > YN > YY > NY
		13.	YN > NY > NN > YY
		14.	YN > NY > YY > NN
		15.	YN > NN > NY > YY
		16.	YN > YY > NY > NN

In the general case, we encounter 2^p issue combinations and $(2p)!$ possible preference orders for p dichotomous issues. Remarkably, the percentage of orders assuming separable preferences already falls below 1% if $p \geq 3$ (Brams et al., 1997, p. 5).[61]

In addition to arguing that NSP are not rare but widespread, a clear definition also has to explain what NSP are not. Nonseparability should not be confused with elementary cause-effect relationships, in which an exogenous alteration causes an adjustment of a related issue. For example, the reduction of a budget leads to less spending. With nonseparability, the restrictions (e.g., the budget constraint) and the resulting adaptation (e.g., the spending decisions) are endogenously determined. They need not necessarily be made by the same actor, but they are part of the same decision-making process.

[59] Bradley et al. (2005) looked for differences in the preference ordering of subjects when alternating between discrete and continuous space. In both cases separability implies for multidimensional alternatives that the preference orderings are independent of the outcomes on all other criteria. The authors argued that in economics and when using spatial models "it is natural to assume that alternative sets are continuous. But in multiple elections they are naturally discrete; in simultaneous referendums, they are binary" (Bradley et al., 2005, p. 335). They stated that in continuous context separability is a very strong property whereas in the discrete case it imposes much weaker restrictions on an ordering. However, common set operations which preserve separability (cf. Gorman's theorems, Gorman, 1968) never apply in discrete space.

[60] This tabular illustration is based on Brams et al. (1997, p. 30, table 1).

[61] 3 dichotomous issues can be combined in 8 possible ways: 111, 110, 101, 100, 011, 010, 001 and 000. These 8 combinations can be arranged in $8! = 40320$ orders of which only 384 can be derived by assuming separable preferences.

2.2 Introductory examples

The following examples clarify NSP further. Moving from hypothetical to empirical cases, I choose short studies verifying the occurrence of NSP in actual "real-world" politics. Please note that these examples serve an illustrative purpose. None of them resembles an all-encompassing case study. The background research has been conducted thoughtfully, but the main goal was to highlight the prevalent nonseparabilities. Thus, I include literature for further reading in all studies.[62] The examples discuss various levels of government, and range from the European over the national and federal to the local level of politics.

The multitude of examples illustrates the commonness of NSP. In addition, they also highlight specific aspects of the nonseparability concept in concrete terms. I start at the European level, presenting a first case documenting political dispute characterized by NSP. This illustrates plainly the multidimensionality of legislative issues, which has to be considered when assessing the corresponding policy decisions. The same holds for the second example, which is located in the national realm of politics. In my report I focus on the privatization debate on Deutsche Bahn, specifically on two sub-questions in order to make their interrelation clear; i.e., the starting point for nonseparability. These first two instances clarify the relevant reasoning and preference dependencies when encountering NSP in the domain of politics. The third example illustrates the consequences of badly coordinated policy decisions. More precisely, the lack of agreement on interrelated issues leads to less than optimal social results, as in the reform process of the German Bundeswehr. This pattern is a very important aspect, as it introduces the interrelation of institutional structure with nonseparability discussed in Sec. 2.3. A prominent part of my study is devoted to the question of how one can assess and measure NSP (cf. Sec. 2.4.6). The final two examples take on this problem. A survey on the future of the German pension system reported (by mistake) NSP for a large share of the German population. Looking more closely at this study, I disclose the problems of a premature assessment. To further strengthen the connection to reality, I investigate an actual referendum conducted in Heidelberg. This local politics example highlights potential problems when looking for interrelations between issues through public polls.

2.2.1 The European Union regulation on chemicals

In its 6^{th} Term (2004-2009) the EP enacted the regulation for registration, evaluation, authorization and restriction of chemicals (COM(2003) 644, further REACH). The regulation sets out the obligations of the industrial sector to assess the risk of chemicals. The industrial sector is obligated to limit the consumption of perilous substances and to substitute them with less hazardous alternatives if possible. Also, consumers must be allowed to

[62] As all studies use secondary data I include also references to the original sources of the data.

access the appropriate information. The changes and improvements made must then be documented for revision by the European chemicals agency (ECHA).[63]

The protocol of the original EP debate on the regulation (which took place at the 17^{th} November 2005) offers a detailed insight into a political dispute characterized by NSP.[64] The discussion of REACH had been highly ideologically driven right from the beginning[65] and remains a controversial issue (DBT, 2012). The ideological dimension in particular made it difficult to reach an agreement. In addition, many single components of the regulation led to one of the most complicated legislative processes ever.[66]

Firstly, the parliament had to decide what chemicals would be included in the regulation. Secondly, it had to decide on how extensive the necessary documentation was to be? The protocols of the parliamentary debate and associated committees state the variety of issues that needed to be considered (e.g., Mann, 2005): environmental protection, consumer protection, competitiveness of industrial sectors, privacy and patent law, sponsorship of research and development, cost estimations, extent of information gathering and disclosure, etc. These issues are interrelated in many ways. They inhibit various trade-offs between different objectives as environmental protection vs. commodity prices, cost absorption by firms vs. competitiveness of the affected industries, disclosure of trade secret vs. information rights of consumers, etc. It is no wonder the regulation led to a lengthy and controversial debate. The final compromise was seen by many MEPs as a very well-balanced scheme. It would achieve a durable "balance on this delicate, complex and controversial decision" (Statement of Lord Bach speaking for the UK presidency, EP, 2005)

This example of a parliamentary debate emphasizes the possible complexity of legislative decision-making. The dependencies led to an incredibly diverse consultation process. Research investigating the field of legislation has to consider NSP in case such multidimensional law proposals are analyzed. It is incorrect to only examine single issues. This would pose a threat to the inference one draws from empirical results, which would be biased, for example, when identifying decisive aspects of the debate or evaluating how important the various aspects were to the different political groups.

2.2.2 The privatization debate on Deutsche Bahn

Another simple example for conditionality of political preferences is the privatization debate on Deutsche Bahn (DB). The DB was founded in 1994 and is the largest rail trans-

[63]For more information on REACH please consult the information archive of the "Directorate-general of the European commission for enterprise and industry" at `http://ec.europa.eu/enterprise/sectors/chemicals/reach/how-it-works/index_en.htm` (accessed September 2, 2012).

[64]The protocol is accessible at `http://www.europarl.europa.eu/sides/getdoc.do?pubref=-//ep//text+im-press+20051116ipr02381+0+doc+xml+v0//en` (accessed September 2, 2012).

[65]Cf. statement in the EP (2005) of Alexander Graf Lambsdorff from the alliance of Liberals and Democrats for Europe (ALDE), MEP of the 6^{th} EP and deputy chairman of the German Free Democratic Party.

[66]Cf. statement in the EP (2005) of Thomas Mann from the European People's Party-European Democrats (EPP-ED), MEP of the 6^{th} EP and REACH Rapporteur of the Employment and Social Committee.

port and railway infrastructure company in Central Europe. It is structured as a group of companies for specific sectors such as, e.g., passenger transport, railway infrastructure and logistics (Radke, 2011). Until today, the Federal Republic of Germany retains all the shares. Since 2006, there has been an ongoing discussion about potentially privatizing the company (Brunnhuber, 2006). The possibilities discussed ranged from a sale of shares of the existing (vertically) integrated group, to a complete sale of the privatized transport companies after the separation of the transport and rail network. Most economic studies called for a separation of railway infrastructure and railway operations. A special report of the Monopolies Commission advocated that the railway infrastructure constitutes a natural monopoly and should therefore remain in state ownership (Basedow et al., 2006). On the other hand, the full privatization of railway transport is economically desirable. Bogart (2009, 2010) and Lalive and Schmutzler (2008) underlined that competition is the decisive factor for growth, quality and reliability in passenger transport by rail. It is exactly this competitive aspect that would be lost if the initial public offer (IPO) were to yield a vertically integrated enterprise. By further retaining the railway network, the DB would possess numerous discriminatory possibilities (Basedow et al., 2006, p. 27). Opponents of the IPO referred to the bad example of railway privatization in the UK and the massive "GB rail efficiency gap" (McNulty, 2011, p. 72).

In 2007, the government of Christian and Social Democrats had agreed on a concept called "holding model": The DB would be preserved as an integrated group, and yet investors would be allowed to participate. Yet, this would only apply to the transport and logistics sector of the companies. However, this plan was abruptly stopped by the global financial crisis (DPA, 2011b). At first this was only meant to be a temporary measure. But it currently appears as if the IPO is postponed into the far future. Tedious technical issues, a lack of reserve vehicles and weather problems have also hampered the IPO intentions (DPA, 2011a).

While the debate has not yet found a satisfactory end, it is still very well suited to clarify NSP. One aspect in the political debate focused on privatization in the strict sense, i.e., the sale of state shares of the corporation through an IPO. This is an idea strongly rejected by the political left-wing (i.e., the parties The Left and The Greens). Others associated with this decision on privatization the question of how to handle a separation of railway infrastructure and railway operations. The Liberals and Christian Democrats supported the IPO of the DB *only if* the separation of infrastructure and transportation would secure sufficient competition.[67] The Social Democrats, however, preferred privatization *only if* the DB would continue to exist as an integrated group.[68]

The point where NSP come into play is marked by the expression "*only if*". So, we observe no NSP for the left-wing parties. They rejected privatization of the DB no matter

[67] Press release of the Liberal parties' presidium (FDP, 2007): "The separation of network and transport is essential."

[68] Cf. the party program of the Social Democrats for the general election 2009.

what. Yet, the phrase "*only if*" makes the conditional binding of the two components very clear for the other parties. While the Social Democrats rejected privatization if the DB were to be split up, they advocated the IPO if the DB were to be preserved as integrated group. Here, the first preference on one issue (privatization) changes depending on the outcome of another issue (separation). This corresponds perfectly with the definition of NSP by Enelow and Hinich (1984, p. 18ff). Finding a solution to this kind of political task is not possible by only focusing on one of the aspects. The next example shows the results if politicians attempt it nevertheless.

2.2.3 The reform of the German Bundeswehr

In early 2010, the German Minister of Defense, Karl-Theodor zu Guttenberg, initiated an analysis to identify the strengths and weaknesses of the German Federal Armed Forces (Bundeswehr, further BW). The study was carried out by a commission, which recommended a comprehensive restructuring of the armed forces. The proposed reform should adapt the defense resources to Germany's current and future security challenges (Weise et al., 2010). In October 2010, zu Guttenberg presented to the public different models of how the future structure of the BW could look like. Two main features dominated the proposed reform: the abolition of the compulsory military service, and the closure of various military sites throughout Germany.

Compulsory military service has been impregnable for a long time in Germany. It was seen as an important link between civil society and the military (Frevert, 2001). Yet the commission concluded that, to meet the current and future tasks of the BW, a more professional and specialized military with highly trained and equipped troops was necessary. This, however, could not be achieved in a compulsory military service of just six months. While this decision is a politically highly interesting topic, the focus of this introductory example lies on the second important feature of the reform, the closure of military sites.

Military bases are frequently found in rural areas. Often, these sites have a long tradition, and are soundly connected with their region nearby. They are an important source for income, tax revenues, job opportunities, etc. If a base is closed the whole region also suffers and loses an important economic pillar (Rudolf, 2011). When soldiers, workers and their families move away, even small retailers, pubs and supermarkets feel the change. In 2011, it was announced that the BW should be reduced from around 250,000 to 180,000 soldiers and that 31 of 400 military sites in Germany would be closed. The affected municipalities feared economic losses up to millions, depopulation and vacancy (Juettner et al., 2011). Therefore, it is not surprising that many federal states and municipalities called for compensation and assistance packages for the abandoned areas (Beck, 2011).

The relationships between local municipalities and the army have always been a give and take. Many communities undertook investment in infrastructure projects to satisfy and retain their military sites. This is precisely where the problem occurred. Many regions had been overconfident that their bases would not be abandoned. Under this presumption of preservation, considerable expenditures have been made; over the last 5 years preceding the reform € 160 million had been spent on infrastructure projects at military sites which are destined to be closed (Bundesministerium der Verteidigung, 2012).[69] Expecting a troop withdrawal, the local authorities would not have made these investments but preferred to consolidate their budgets.

One important but simple example is the 5,000-resident city Kusel. The military site near the city is one of five sites in the Rhineland-Palatinate which will be closed. Nearly 1,200 soldiers will move away. The decision hit the local policy makers by surprise. Especially, as in recent considerations of the army reform even strengthening the site had been discussed. At the moment of the decision, civil investments on the military site took place, a new gymnasium and administrative buildings were being built (Juettner et al., 2011).

A reason the coordination failed might be the varying distribution of preferences and conditionalities. On the one hand, the municipalities had NSP about infrastructure expenditures and the continuation of the facilities (investment only on condition of preserved existence). On the other hand, the federal government was probably not at all interested in infrastructure expenses, as the reform in itself was already highly complicated and demanding (future troop size and composition, restructuring due to the abolition of the compulsory military service).

This example shows the difficulties that arise when decision-making is divided on problems which are related to each other. The lack of coordination between infrastructure investments at the local and defense policy reforms on the state level led to avoidable cost. It is not necessarily the case that the decision-making bodies are situated at different levels of government. It could also happen between two different ministries or two neighboring cities which do not coordinate. So, they end up which less than optimal outcomes, like investments in the infrastructure at locations shortly to be abandoned.

2.2.4 The future of the German pension system

I have already argued a few times that NSP are often ignored although they are actually commonplace. Ironically, NSP are sometimes identified in instances in which they do not really exist. This is due to a misunderstanding of their basic principle. Based on a current survey study I clarify this misconception. This instance is an example of how not to deal with NSP.

On August 29th and 30th 2012, the survey institute TNS-Emnid surveyed 1001 German citizens for their views on the future of the German pension system. The results caught the attention of the media quickly, as the survey followed a recent and broad discussion on the peril of old age poverty. The debate started after a statement of the German Federal

[69] Even though this may be a relatively small amount compared to the overall € 31.55 billion in the German defense budget (Deutscher Bundestag, 2011), this is an immense sum of money.

Minister of Labour and Social Affairs, Ursula von der Leyen, on a prospective lowering of the pension level to around 43% of the depositor's nominal wage (von der Leyen, 2012).

Media reports on the survey said that a majority of 51% of Germans citizens was willing to accept the proposed lowering. Only 39% were in favor of increasing the pension contributions in order to prevent the lower pension level. At the same time, 32% of the respondents were in favor of supporting particularly low pensions with taxes, in return for lowering the contributions to the pension fund. Another 19% stated that future retirees should make provisions for themselves by investing in private pension insurance.[70]

I do not question the reported numbers, but I do point out a certain aspect of nonseparability in the coverage of the survey. In (nearly) all news reports it was stated that the lowering was accepted *in return* for the tax means subsidies (e.g., Christians, 2012). This implies a conditional relationship, as the lowering is conditioned on the subsidies. The two simple but separate questions "Do you support or reject the lowering of the pension level?" and "Do you support or reject to subsidize particularly low pensions?" are not suited to identify such a trade-off. They only tell us just how many respondents support a lower pension level, and how many people support tax subsidies.

No kind of conditionality or trade-off can be identified in this way. I do not argue that such a trade-off is absent; in fact, I honestly do not know, and would probably also agree that an interrelation is likely (to a certain degree). The point I wish to make is that these questions cannot tell us whether or not the German public has such considerations. A correct survey question, aiming to identify a possible nonseparability between these issues, would read as "Assuming that particularly low pensions are increased with tax means, do you support or reject the lowering of the pension level?"

This example shows that NSP and surveys have a "special" relationship. On the one hand, surveys possess the ability to identify even highly complex preference relations. This holds at least in theory, and it has yet to be proven empirically. Often, nonseparability is avoided because it takes time and entails complex requirements. On the other hand, not accounting for NSP in surveys risks mistaking a respondent's conditional response for the genuine preference, and vice versa (Lacy, 2001a). In contrast to this example, Sec. 2.4.6 clarifies how a survey investigates NSP properly.

2.2.5 The construction of a new convention center in Heidelberg

This example looks at a referendum at the local level of politics. In 2010, the citizens of Heidelberg were called to vote in a referendum on the construction of a new convention center, respectively, the extension of the existing city hall with an annex to serve as con-

[70]For a more comprehensive review of the survey and the public discussion on the reform of the German pension system in general please consult the Focus money online dossier at `http://www.focus.de/schlagwoerter/themen/r/rentenniveau` (accessed December 11, 2012).

vention center.[71] Remarkably, the referendum was due because more than 18,000 citizens signed a petition that called for a public vote on the extension of the city hall. The referendum then took place on 25^{th} July 2010 and with a two-thirds majority (by a turnout of 38.9%) the extension was rejected.[72]

As often in such cases, this public vote preceded a lengthy and difficult discussion. Plans to build a new convention center near the Heidelberg central station have existed since the 1990s. One after another, four official tenders (1996, 2000, 2004 and 2006) provided no viable concept. After all plans had failed, including those involving private investors, the Heidelberg City Council decided in 2008 to change the location. Instead of building a new convention center near the central station, it decided to modify the existing city hall by an extension. The two locations are about 3 km away from each other. The existing city hall was built from 1901 to 1903 and its Renaissance architecture fits very well into the historic old town of Heidelberg. Yet, its capacity (max. 3500 people) and alignment could no longer keep up with the demands of modern convention centers.[73]

The issues in the debate were controversial. Proponents of the expansion highlighted the fact that a conference center in the old town would be a unique selling point for Heidelberg. In many cities, such centers are located near train stations or outside the inner cities. A building located in the historic old town would result in a big boost to meetings, conferences and tourism in general. In addition, an extension would also be much less costly than constructing of a new building from scratch at a different location.[74] Opponents of the expansion pointed to the increase in tourism, too. Heidelberg's old town, they claimed, would not be able to cope with more tourists. Additional tourism and traffic would be harmful and difficult to manage because of the poor urban transport links. This is especially true in comparison to the location near the train station and its optimal infrastructure connections. So far, this discussion seems like a simple one-dimensional question. The key reason why this referendum is mentioned here is the following: the opponents had a second objection, the architectural design of the extension. They claimed it would distort the historic facade of the town hall and disfigure the homogeneous impressions and decades-old skyline of the old town center.[75]

The referendum consisted of a single question, which read as "Should the city of Heidelberg build an addition or new building for an expanded convention center at the location

[71] Homepage of the architectural design competition: `http://www.heidelberg.de/servlet/PB/menu/1198771/index.html` (accessed September 19, 2012).

[72] Official election archive of the city of Heidelberg: `http://www.heidelberg.de/servlet/PB/menu/1208720/index.html` (accessed September 27, 2012).

[73] The City Journal Heidelberg provided a detailed online dossier on the referendum at `http://ww2.heidelberg.de/stadtblatt-online/index.php?artikel_id=3744&bf=` (accessed September 20, 2012). It is still accessible and contains information on the extension plans, the different positions of the local parties in the city council, statements from various citizen groups, etc. The city administration also offers a comprehensive summary of the events at `http://www.heidelberg.de/servlet/PB/menu/1125805/index.html` (accessed January 7, 2013).

[74] Cf. representative for the proponents the then acting mayor Würzner (2010).

[75] Cf. representative for the opponents Lask (2010).

of the city hall?" The possible answers were 'yes' and 'no'. The architecture of the building was not part of the vote, even though it had just been the design of the new structure which had brought a whole new level of controversy into the discussion (cf. Möslinger, 2010). The referendum contained no possibility for people who wanted a convention center at this location with a different design to express their preferences unambiguously. If participants chose "yes", they approved location *and* design. If they chose "no", they rejected *both*. Alternative designs, an important part of the public debate, were not part of the referendum.

Again, this points out that holding a referendum alone is not enough. To increase democratic legitimacy the details of the implementation must also be considered (Lacy and Niou, 2000). This problem is in line with the challenges Lacy and Niou (1994, 2000) identified for referendums. They argued that if the topic in question contains interrelated issues, a single dichotomous vote may not be enough to enable all participants to express their preferences.

2.3 Do delegation, decentralization and specialization help?

The previous sections demonstrate that the concept of NSP is linked to complex decision situations. Multidimensionality is obvious, but the specific dependencies between the issues contributes to the intricacy as well. Even collective decision-making on simple issues can be a costly, i.e., time-consuming and nerve-racking, process; let alone reaching an agreement on complex policy proposals.[76] This raises the question of whether one can reduce this complexity and circumvent the requirements of dealing with nonseparability.

THE INSTRUMENTS OF ORGANIZATIONAL THEORY

The ubiquitous remedies to reduce complexity and inefficiency found in organizational theory are delegation, decentralization and specialization.[77] Kaiser (2007b, p. 128) points out that representative democracies use the delegation of powers of attorney a lot to enable collective action at all.[78] Yet, highly decentralized and specialized policy-making processes must be accompanied by some powerful central coordination or regulatory authority.[79] Otherwise decisions are made under uncertainty, bearing the risk of less than optimal and inefficient outcomes.

The main purposes of delegating powers in the firm, the bureaucracy or the political system at large are the reduction of transaction costs (e.g., Epstein and O'Halloran, 1999) and information retrieval (e.g., Chongwoo and In-Uck, 2011; Gautier and Paolini, 2007).

[76] Already Simon (1962) proved that complexity occurs in a large array of topics as, e.g., social, biological, physical and symbolic systems.

[77] These organizational tools correspond well with the idea of "polycentric systems" in solving collective-action problems (cf. Ostrom, 1999, 2008).

[78] Cf. Strom et al. (2003) for an overview on delegation patterns in western parliamentary democracies.

[79] For a formal model and empirical test of administrative procedures cf. Epstein and O'Halloran (1996).

Departments and committees are manned by specialists who know the expected effects of competing policy alternatives as well as the key players in their area of expertise (e.g., Dewan and Hortala-Vallve, 2009; Gilligan and Krehbiel, 1990). Hence, delegation and specialization supposedly lead to faster decision processes and superior policies. Decentralization is justified by similar arguments and applied when the locals are believed to hold superior information on the effects of competing policies. The question concerning the optimal vertical and horizontal division of competences can be found in the literature on federalism (e.g., Buchanan and Tullock, 1962; Oates, 2005), the theory of the firm (e.g., Tirole, 1988; Williamson, 1975)[80], the division of labor within government and parliament (e.g., Gailmard, 2009; Gilligan and Krehbiel, 1987) and intergovernmental as well as international agreements (e.g., Koremenos, 2008).[81]

Delegation and decentralization bear potential problems. Firstly, the specialized agents may follow their own policy agenda and misrepresent the preferences of their principals (e.g., the voters). The danger of bureaucratic drift increases with information asymmetry and conflicting interests between principals and agents (e.g., McCubbins and Schwartz, 1984; McCubbins et al., 1987). Secondly, delegation in general and decentralization in particular may forestall economies of scale and scope (e.g., Weingast, 2009). With respect to the production of policies, the former is prevalent in the case of decentralization: Why should each federal state have its own army? Why should each county regulate solar thermal heating systems individually? By contrast, economies of scope are more of an issue with respect to the horizontal allocation of competences: Why should the environmental authorities, who know all the effects of existing industrial filters, leave the regulation of the respective sector to the Ministry of Industry?

Under the assumption of separable preferences, the application of delegation, decentralization and specialization has proven to reduce costs (Epstein and O'Halloran, 1999) and complexity effectively (e.g., Severinov, 2008; Tsebelis, 1994). They are approved and often used organizational measures. Yet, this tells us nothing about the underlying driving forces of complex decision-making. Each measure causes the separation of decision-making with respect to either a spatial, temporal, hierarchical or administrative dimension (cf. Bendor et al., 2000; Bendor and Meirowitz, 2004). This contradicts the basic tenet of nonseparability stating that interlinked parts should not be separated. The small case study on the reform process of the German Bundeswehr (cf. Sec. 2.2.3) provides a first insight into what happens otherwise. It clearly describes the consequences of badly coordinated decisions. In addition to this specific sample, I am looking for general patterns concerning institutional structures and nonseparability. The question is, e.g., what happens if single decisions of a problem affected by NSP are delegated.

[80] Cf. Grant (1996) for a comprehensive knowledge-based theory of the firm. He looked particular into hierarchy and the distribution of decision-making authority.

[81] Koremenos (2008) investigated "when, what, and why do states choose to delegate". She found that "the presence of a complex problem increases the probability of external delegation [to a third party outside of the agreement] by nearly forty percent and internal delegation [to a collective formed by the members of the agreement themselves] by twenty-one percent" (Koremenos, 2008, p. 169).

THE INTERPLAY OF INSTITUTIONS AND PREFERENCES

Recently, behavioral economists started to scrutinize the effect of the institutional structure (e.g., Ambrose and Schminke, 2003; Wilson and Eckel, 2011) on the motivations of individual choices (e.g., Bandiera et al., 2005; Bolton et al., 2005). In this context, institutions comprise decision-making rules, proceedings as well as algorithms (Kröning and Strichman, 2008). The investigations are not solely restricted to simple economic games. Ehrhart et al. (2007) examined the impact on the setting-up of a budget, a common policy issue. The results are not decisive but point to the relevance of the decision-making environment as "preferences may be sensitive to the choice process" (Sen, 1997, p. 745) itself. This is of central interest for research in the field of empirical institutional analysis, which seeks to determine these "causal mechanisms, i.e., the way how institutions affect behavior structuring" (Kaiser, 2007b, p. 120).[82]

My research is based on the interplay of decision-making procedures and preferences of the decision-makers. Already Plott (1979, 1991) argued that the interaction of preferences and institutions determines policy outcomes. The preferences are channeled through institutions which condition actors' behavior and combine them to collective choices. A change in either preferences or institutions, while the other remains constant, might change outcomes, but need not necessarily do so. This allegory is referred to as the "fundamental equation of politics" (Hinich and Munger, 1997, p. 17) and depicted in Eqn. 2.1.[83]

$$preferences \circledast institution = outcomes \tag{2.1}$$

The equation is rather self-explanatory. Outcomes merely refer to the result of a decision-making process, whether it is what to cook for dinner, which movie to go to, or whether or not to build a new city hall. In other words, this means the societal consequences of decisions made (Lane and Ersson, 2000). Preferences are straightforward too, as they resemble what a person prefers in a given set of choices (Shepsle and Bonchek, 1997). The most simple form is "given by a binary relation over the set of options" (Barrett and Salles, 2006, p. 1). People may differ in their preferences in terms of direction, salience and over time (Ehmke et al., 2005). Complementary concepts as, e.g., endogenous preferences can also be incorporated (cf. Grendstad and Selle, 1995). This supplements the

[82] For empirical examples cf. Diermeier and Gailmard (2006) who showed that social preferences depend on the decision-making process, or Sobel (2005, p. 392), who summarized "context-dependent preferences" as to permit "the strategic context to determine the nature of individual preferences."

[83] The importance of this fundamental principle is aptly illustrated by the blog "Rule 22" maintained by Ragusa et al. (2012). The name refers to a standing rule of the United States Senate most commonly associated with the filibuster. The authors are motivated by "the intention to forward the insights into the public debate" and analyze political events through a general institutional perspective. The blog entry "Things Institutionalists Know that You Should: Plott's Equation" (`http://rule22.wordpress.com/2011/09/22/things-institutionalists-know-that-you-should-plotts-equation`, posted on September 22, 2011 by Nate Birkhead) emphasized the fundamental role of Plott's equation along various examples.

standard approach, which assumes "that people harbor a stable, well-defined, and discernible order of preferences" (Simon et al., 2004, p. 331).

The term most difficult to circumscribe is "institution". Of course, the word itself is common-place and often refers to any formal structure or organization. Yet, according to North (1990, p. 79ff), a distinction between institution and organization is pertinent. North classified organizations as the endogenous best response of humans to their institutional environment. On the other hand, he defined institutions as "the constraints that humans devise to shape human interaction" (North, 1990, p. 3). Thus, "institutions are the rules of the game" (North, 1997); they are set first. Within these rules, an actor's behavior creates organizations.

A central question is where these rules come from. In this context, Hall and Taylor (1996) provided an excellent overview of the different research paradigms which have appeared within the "new institutionalism": rational choice, historical and sociological. Rational choice institutionalism has developed the most "precise conception of the relationship between institutions and behavior" (Hall and Taylor, 1996, p. 950). Yet, this results from a highly simplified image of man (Opp, 2004). It has provided major scientific contributions to political research, but has only limited explanation power in explaining the emergence of institutions (Hall and Taylor, 1996, p. 952). Historical and sociological institutionalism provides more insights through their focus on existing power relations and how they shape further developments. In sum, Hall and Taylor (1996) do not argue in favor of one approach but rather "for greater interchange among them" (Hall and Taylor, 1996, p. 955). This implies for the emergence of institutions that they are inherited, or else developed, by actors when the need arises. This fits with North (1990, p. 3) and his argument for "humanly devised" institutions.

Ostrom (1986) also focused on the multiple usage of the term institution. Reformulating Eqn. 2.1 to an equation system as in Eqn. 2.2, she argued that we "need to address questions concerning the origin and change of rule configurations in use. How do individuals evolve a particular rule configuration? What factors affect the likelihood of their following a set of rules? What affects the enforcement of rules?" (Ostrom, 1986, p. 22)

$$\begin{array}{ccccccc} \textit{action structure} & \circledast & \textit{decision model} & & & = & \textit{outcomes} \\ \textit{rules} & \circledast & \textit{physical laws} & \circledast & \textit{behavioral laws} & = & \textit{action structure} \end{array} \tag{2.2}$$

In Eqn. 2.2 rules (i.e., institutions) determine, in combination with physical and behavioral laws, the feasible action structure of actors. The interaction of this structure with the decision model then produces the resulting outcomes. An example of such an extended system was provided by Lane (2000). He argued that the most important action structure is "the market", which offers rules for interaction of economic and political interests.

A broad strand of literature agrees that interests and institutions affect social outcomes. Krehbiel (1986, p. 544) considered political outcomes as products of decision-makers' preferences and institutional features. Looking into legislative committees he described committee strategies as "determined by preferences and institutions" (Krehbiel, 1986, p. 555). Current contributions also refer to this research framework. Thomson et al. (2006, p. 9) classified it as "one of the major preoccupations of modern political science: the interplay between institutions and preferences in determining policy outcomes." Schofield (2008, p. 2) stated that social choice theory in its very core "seeks to understand the connection between individual preferences, institutional rules and outcomes."

An unsettled matter concerns the relative importance of preferences and institutions for the resulting outcomes. Dowding and King (1995, p. 7) advocated the dominant role of institutions as, "generally speaking, the institutions of politics provide a larger part of the explanation than do preferences". To make such a strong and general claim is controversial, as it extends to a large variety of decision problems and contexts. Yet, "most of the new institutionalism still leaves a role for preferences to play" (Stoll, 2013, p. 7) and prefers to emphasize the importance of institutions in relation to preferences. Here, Riker (1980, p. 20) predicated strongly that social research "cannot study simply tastes and values, but must study institutions as well" and defined institutions as "rules about behavior, especially about making decisions" (Riker, 1982b, p. 4).

Various previous contributions have focused on the influence of procedural rules and demonstrated their stabilizing effect (e.g., Shepsle, 1979). Without institutions, any multidimensional decision problem is inherently unstable (McKelvey, 1976). The field of comparative politics features a vast literature on the variance in institutional arrangements across countries. Well known fields of research are regime types, electoral and party systems, the structures of government, civil-military relations, corporatism, social cleavages, etc. (for an overview cf. Lijphart, 1971, 1999). Only one prominent example in the field of analytical politics is the so called "structure-induced equilibria" (Shepsle and Weingast, 2004), which takes the structuring capabilities of institutions as given.

In recent studies, "institutions are more and more regarded as result (or equilibrium solution) from strategic action" (Kaiser, 2007b, p. 124).[84] This highlights that "preferences matter, too" (Stoll, 2004, p. 4). Turning the focus back from the institutional structure alone, Lane (2000) argued that interests are important, and that institutionalism overstates the role of procedural rules. Stoll (2004) reviewed the ways in which preferences interact with political institutions to shape party systems. Overall, she acknowledged that "political outcomes are a function of both political institutions [...] and the preferences of the citizenry" (Stoll, 2004, p. 1). Yet, she emphasized that "institutions do not - and cannot - tell the whole story: preferences have work to do as well." (Stoll, 2004, p. 2)

As with institutions, views on preferences changed. The current state of research considers preferences not as exogenously fixed, as they were in classic rational choice theory

[84] When assessing the degree of risk sharing and redistribution in different federal fiscal constitutions, Persson and Tabellini (1992) aptly illustrated how different forms of interaction (e.g., voting or bargaining) lead to different institutional frameworks.

(Simon et al., 2004). They "rather are reconstructed in the course of decision making" (Simon et al., 2004, p. 335). March and Olsen (2006, p. 689) emphasized that rules are followed if they are judged "appropriate [...] in a specific type of situation." To incorporate these insights in Eqn. 2.1 leads to Eqn. 2.3. The outcome is still defined by preferences and institutions. However, it also accounts for the possibility of a "feedback loop" between preferences and institutions.[85]

$$\left(preferences \underset{\leftarrow}{\overset{\rightarrow}{}} institution\right) = outcome \tag{2.3}$$

Holyoak and Simon (1999) found that in legal decision-making shifts in one task can trigger according alterations in subsequent tasks involving similar underlying issues. In addition to the pure existence of such shifts, it is important to search for regular patterns. Looking into legislative bargaining, Miller and Vanberg (2014, p. 18) gathered evidence indicating that the "unanimity rule motivates subjects to be more 'bullish' in their bargaining behavior". This is no single or isolated finding; performing a literature review on the effects of procedure on social interaction, Mertins (2008, p. 10) identified that process fairness affects fairness perceptions. She concluded that "clear evidence for procedures in influencing human decision-making exists" (Mertins, 2008, p. 31). Following this insight, the current state of research moves away from classic but static approaches (Wandling, 2012).[86]

Including dynamics is complex (and difficult to operationalize) but closer to reality; such a dynamic is found when variations of the procedural rule lead to a change in preferences. Nonseparability adds a new element to the interplay of institutions and preferences. The separation of decision-making competences, or defining one part of a decision, causes an adjustment of the remaining (conditional) preferences.[87] Any organizational measure, whether it is delegation, decentralization or specialization, implies a multilevel or multistage decision (Strom, 1990, p. 107ff). Thus, it causes the separation of decisions belonging to one overarching question. Yet, this violates the basic tenet of separability, stating that interlinked parts should not be separated. This emphasizes again the commonness of NSP, and highlights that delegation, decentralization or specialization are no solutions to, but rather part of, the unanswered (empirical) question about the influence of nonseparability.

The next section looks into the theoretical aspects of nonseparability and focuses on the question of how to operationalize NSP. In the subsequent part, I use the framework outlined in compound with the argument that NSP are highly relevant in the context of

[85] This is different from classical multiattribute decision theory (e.g., Edwards and Newman, 1982; Keeney and Raiffa, 1976) as preferences cannot be added up or weighted accordingly to specific environments.

[86] Lichbach (2001) juxtaposed rationalist, culturalist, and structuralist research approaches against each other. The comparison clarified that these approaches cannot go their traditionally separate ways anymore. They must submit themselves to the mutual exchange of criticism in order to remain valuable.

[87] Cf. Sengul et al. (2012) for a discussion on the strategic aspect of delegation.

organizational structures and explicate common arguments on the effect of NSP with respect to organizational measures of delegating and sequencing. Those are not completely new and have existed in the theoretical literature for quite some time, but their empirical verification is still pending.

2.4 Theoretical foundations

This section discusses the theoretical foundations for analyzing NSP. As the concept is deep-seated in the theoretical literature (for an overview cf. Strom, 1990), it combines existing knowledge of different fields into one comprehensive discourse. I start the theoretical discussion in Sec. 2.4.1 by stating basic definitions. Next, I discuss in Sec. 2.4.2 the usage and operation of spatial models in analytical politics. Spatial theory serves as main methodological tool for analyzing preferences in general (Shepsle and Bonchek, 1997). Following this approach, I illustrate in Sec. 2.4.3 the implementation of NSP into a simple spatial model. Sec. 2.4.4 discusses briefly the magnitude of NSP in such a model, and Sec. 2.4.5 introduces a specific extension to the operationalization with respect to the realm of politics. Finally, as every model evaluation must rest on empirical data, Sec. 2.4.6 exemplifies how NSP can be measured.

2.4.1 Basic definitions

Collective decisions on conditional topics are the subject of my research. I have, so far, already talked a lot about '*conditionality*', but skipped a detailed discussion of the term '*collective*'. A collective decision obviously refers to a *group* of individuals. In general, such a group is not limited to any size. However, my laboratory experiment will be restricted to small groups because of the space provided in the lab. Political decision-making to a large extent takes place in committees. I regard a committee as a small to medium-sized group of people who have institutionalized interactions. Following the work of Black (1958, 1991), and in the tradition of rational choice theory, the focus of this study lies on voting procedures and results in such committees.[88]

While it is relatively easy to define an individual decision, this is a more complex endeavor in the case of a collective decision. In this study, a collective decision represents the coordination of actions within a group of relevant actors. More precisely, each participant of the decision-making group chooses one option out of a set of options, and rejects all others (Pritzlaff, 2006, p. 208). The decisions of every involved individual engage with each other on the collective level and form the collective decision of the group.

[88] Cf. Nullmeier and Pritzlaff (2009) for an alternative approach and a more specific look into the concrete deliberation of committee meetings and their internal decision processes.

2.4.2 Using spatial models

A key component when analyzing preferences in the rational choice tradition are *spatial models*. Their underlying assumption states geometric dependence among modeled objects. Fundamental to spatial models is the work of Hotelling (1929), Lerner and Singer (1937) and Smithies (1941). While these authors focused on economics and the competition between companies, the contributions of Downs (1957) and Black (1958) established spatial models in analytical political science, and Davis and Hinich (1966, 1967, 1968) laid the groundwork for spatial theory of voting.[89] Recently, increased efforts and progress have been made to operationalize spatial models of legislative politics (e.g., Benoit and Laver, 2007c; Laver and Schofield, 1998; Slapin and Proksch, 2008).

To operationalize spatial models in politics, one assumes that political agents have rational preferences over a predetermined decision space (cf. Schofield, 2008). Thus, the actors have a clear and unambiguous preference for the realization of each dimension of the space. The point in space where the first preferences for every dimension intersect is called the *ideal position* (IP) of an actor. Each possible individual policy can be represented as point in the decision space (Humphreys and Laver, 2010). For each such point the utility level U of every actor i can be determined as the distance between this point and their IP; i.e., how far for a specific policy is away from the most preferred option.

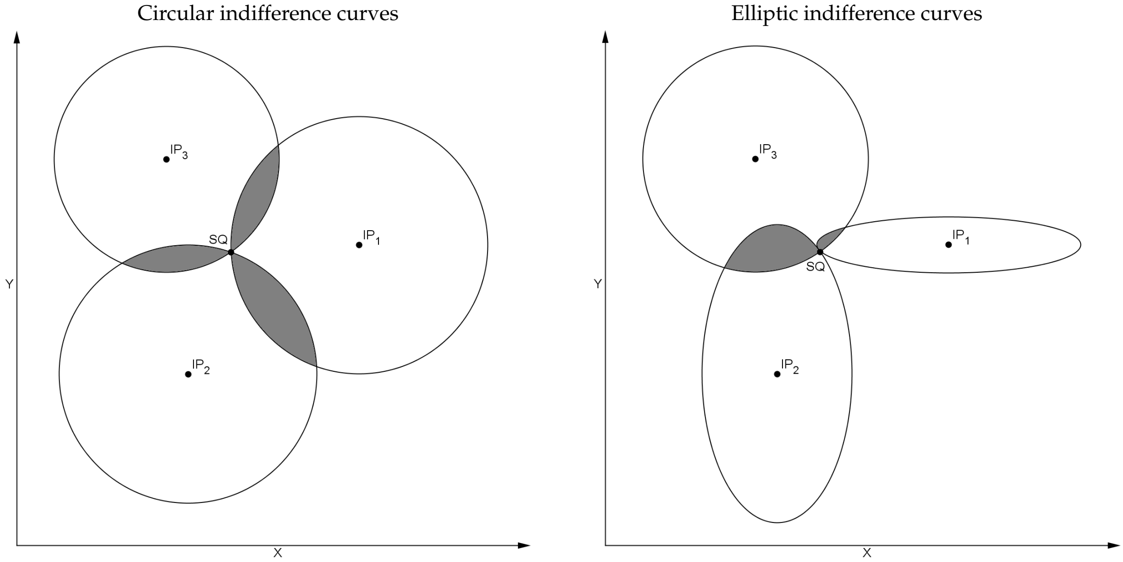
Fig. 2.1 shows a two-dimensional policy space (X,Y) including the IP of three actors and the status quo (SQ) as the currently enacted policy. Every policy closer to an actor's IP raises the utility of this actor and is therefore preferred to the SQ. Colomer (1999) called this the "preferred-to-set concept". It comprises all possible policy outcomes an actor prefers over the current SQ. The left figure shows cycles drawn from each actor's IP through the SQ. These cycles represent the actor's *indifference curves*. At every point on the curve the utility level of the respective actor is the same, thus they are indifferent when choosing between points on the same curve. The form of circles is the result of my assumption in the left figure of both dimensions being equally important to all actors. If one dimension were to be more important to an actor the curve would take the form of an ellipsis. This happens in the figure to the right. A differences in salience of the issues stretches the indifference curves in the dimension of the lesser important issue. This means that for actor 1 the issue Y is more important, and, for actor 2, X outranks Y. Independent of any salience, every point within the curve exhibits a higher and every point outside the curve a lower utility. In the two-dimensional policy space exist infinitely many indifference curves, one passes through each possible combination of X and Y.

The darker areas in Fig. 2.1 resemble policies (X,Y) which two of the three actors would prefer to the SQ, i.e., the indifference curves of two individuals overlap. These areas are called *winset* as they allow a majority of actors to improve from the current situation

[89] Cf. Davis et al. (1970) for a non-technical introduction to (spatial) models of social choice.

Figure 2.1: A simple spatial model

EXPLANATORY NOTE
The figures show a two-dimensional policy space (X,Y), the SQ and IP of three actors. The IP of an actor represents her most preferred policy (or point of the policy-space), which is given by the combination of the two dimensions. Cycles or ellipses around the IP indicate an actor's indifference curves.



(assuming simple majority is the necessary threshold). The winset of the SQ is the set of outcomes that can defeat the SQ in a pairwise comparison, i.e., the set of policies that can replace the existing one.

A related concept is the concept of the *core*. A core is the set of points with an empty winset; i.e., the points that cannot be outdone by any other point under the corresponding decision rule (Tsebelis and Garrett, 2001, p. 21ff). When looking at collective decisions it is important that the necessary majority threshold is reached. In Fig. 2.1 no policy exists for which all three actors are better off, i.e., no winset exists under unanimity rule. Thus, whether the actors will deviate from the current SQ or not depends on the majority necessary for altering the enacted policy. Importantly, the difference in salience in the right figure influences the winset. Here actors 1 and 2 have no common ground for an improvement from the SQ.

This example is formalized as follows: For each point in a two-dimensional decision space (X,Y), the utility level of every actor i can be determined as the utility at their ideal position (U_{IP}) reduced for the distance of the policy in question to it. Accounting for the dimension-specific differences in salience this results in

$$U_i\left(x,y\right) = U_{IP,i} - \left[w_{x,i} \times \left(IP_{x,i} - X\right)^2 + w_{y,i} \times \left(IP_{y,i} - Y\right)^2\right]^{\frac{1}{2}} \tag{2.4}$$

In Eqn. 2.4, the distance between policy and IP in each dimension is weighted with an actor-specific weight w_i. The underlying logic is that if one issue is very important to an actor, then a deviation from their first preference in this dimension has a stronger utility-reducing effect. Eqn. 2.4 makes one additional (and fundamental) assumption I have not discussed so far. The equation represents the distances in terms of *Euclidean space*. The Euclidean metric determines the distance between two points as the root of the sum of squares of differences of the individual dimensions (Hinich and Munger, 1997, p. 76ff).

Before applying these concepts to the case of NSP there are some remarks to be made. Using utility in political science (like many other empirical tools) can be attributed to influence from economics (Webster and Sell, 2007). Thus, it was not developed for "political" questions and should henceforth be scrutinized for its suitability. The main application for utility space in economics is dealing with the allocation or distribution of goods: to identify goods as substitutes or complements, to investigate their (marginal) exchange rate, etc. In politics we mostly deal with policy programs instead of goods. Nevertheless, the use of spatial models and Euclidean space is standard practice in political science. The work of Tsebelis (2002) is just one prominent example.

In the literature, nonseparability has been prominently discussed with respect to spending preferences over public policies which are subject to the same budget (Hinich and Munger, 1997, p. 60ff). Given an overall budget constraint, the amount of money allocated to one policy program influences the spending preferences for all remaining programs. Milyo (2000a,b) showed that the Euclidean metric cannot correctly represent actors' utility functions over such a multidimensional budget-rivalry problem. According to Benoit and Laver (2007c, p. 31), this finding further supported those who criticize the convention of using the Euclidean metric to model political decision-making. Humphreys and Laver (2010, p. 14) argued that "metric assumptions for models of policy-based political decision making are under-researched, and which distance metric - if any - is appropriate for modeling human political preferences remains an open question." Subsequently, I take this criticism very seriously but do not deviate from the conventional Euclidean space.[90]

The main reason I use Euclidean space is the perception of differences by humans. Shepard (1991) studied the perception of (dis)similarity by individuals in a multidimensional setting using the "Minkowski r-metric" in an experiment. In terms of this metric a distance d between the values i and j of dimension x is calculated as

$$d_{ij} = \left[\sum_k \left| x_{ik} - x_{jk} \right|^r \right]^{\frac{1}{r}} \tag{2.5}$$

In his analysis Shepard (1991) iteratively changed r. His goal was to identify the best possible match between the perception of distance by the subjects (determined from their behavior during the experiment) and the distance measure based on the metric. It is noteworthy that the two most prominent metrics can be represented by Eqn. 2.5. The

[90] Only few experimental studies apply other metrics; and if they do, they do not usually find any differences due to the specific metric used (e.g., Berl et al., 1976).

city-block metric[91] for $r = 1$ and the Euclidean metric for $r = 2$. Looking for the most appropriate metric for different choice problems Shepard (1991) found the smallest deviation for the city-block metric using separable dimensions. On the contrary, the Euclidean metric produced its best predictions dealing with nonseparable dimensions.[92] Thus, Euclidean space represents the appropriate metric for my investigation of nonseparability.[93]

2.4.3 How to operationalize nonseparable preferences

By using two-dimensional spatial utility functions, NSP can nicely be illustrated (cf. Hinich and Munger, 1997). The canonical equation for a spatial loss utility function or weighted Euclidean distance (WED) in a d-dimensional policy space is

$$WED\left(\theta, x\right) = \sqrt{\left(\theta - x\right)^T A \left(\theta - x\right)} \tag{2.6}$$

in which θ=(θ_1,θ_2,...,θ_d) describes an actor's unconditional ideal point, x=(x_1,x_2,...,x_d) describes the policy in question and $A = \begin{bmatrix} a_{11} & ... & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & ... & a_{dd} \end{bmatrix}$ describes the importance the actor attaches i) to each dimension (main diagonal) and ii) to the interaction between the dimensions. Thus, the utility of an individual is determined by the weighted distance between an actor's ideal point and the enacted policy.[94]

If we assume that actors preferences are separable across dimension, it holds that $a_{ij} = 0$ unless $i = j$. In other words, a person's "preferences on every issue and set of issues are independent of - or, can be separated from - the outcomes of other issues" (Lacy, 2001a, p. 132). This leads to a utility function represented by a simple "additive model" (Strom, 1990, p. 57) in which both dimensions can be evaluated individually. The deviations from an actor's ideal position can be calculated separately for each dimension and then summed up in the end. Going back to the culinary example of the two couples introduced above it implies that an actor's preference ordering over the disposable wines is independent of the meal and vice versa. This leads to Eqn. 2.7, which elucidates that the pleasure of food and drink is each separately formed for themselves and only assembled into one expression in the end (cf. Eqn. 2.4).

$$WED\left(\theta, x\right) = \sqrt{a_{drinks}\left(\theta_{drinks} - x_{drinks}\right)^2 + a_{food}\left(\theta_{food} - x_{food}\right)^2} \tag{2.7}$$

[91] The city-block metric determines the distance between points as the sum of the absolute differences of the individual dimensions (Torgerson, 1952, 1967, p. 38ff).

[92] Interestingly, Soto and Wasserman (2010) reached the same conclusion when looking at animal behavior.

[93] Most literature looking into the perception of separable dimensions finds that city-block metric provided a far better description of psychological distance relationships (for an overview cf. Lockhead and Pomerantz, 1991). However, allowing for two-dimensional responses of subjects Nosofsky (1985) found that the Euclidean metric provided a far better description of psychological distance relationships.

[94] This complex mathematical equation rests on the same basic principles discussed in Sec. 2.4.2 and Fig. 2.1.

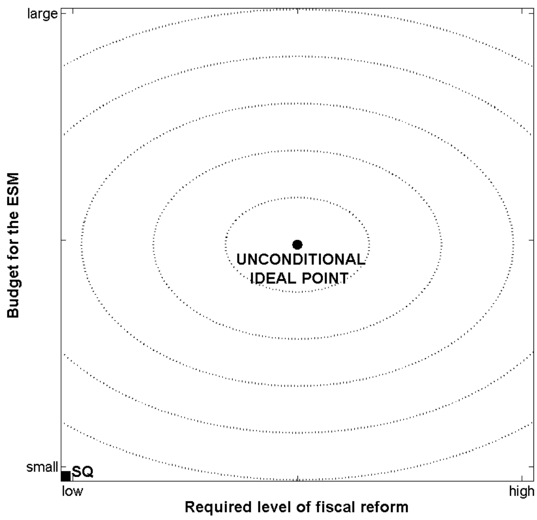
In Eqn. 2.7, θ describes the initial (unconditional) first preferences of an actor concerning either food or drinks, a depicts their respective salience and x the actually chosen selection. With this information it is possible to calculate an actors' "culinary utility" when having dinner. Apparently the pleasure is very dependent on the difference between initial preference and current selection.

Moving closer to a political environment, Fig. 2.2 illustrates the spatial representation of a decision problem for a member of the Euro-group on two separable issues. The member is voting on the budget for the European stability mechanism (ESM) and the required level of fiscal reform the Euro-group demands for its assistance.

Figure 2.2: Indifference curves for separable preferences

EXPLANATORY NOTE

The figure depicts the unconditional ideal point and corresponding indifference curves for a member of the Euro-group who prefers a medium high budget for the ESM and medium strict fiscal requirements. The salience of issue 1 (budget for the ESM) equals 0.7 ($a_{11} = 0.7$) and the salience of issue 2 (required level of fiscal reform) equals 0.3 ($a_{22} = 0.3$). As one issue is more important, the indifference curves have the form of an ellipsis. I assume separable preferences by setting $a_{12} = a_{21} = 0$. The winset originating from the SQ covers the complete issue space.



In contrast, with conditionality the utility in one dimension depends on the outcome of the second dimension. Any $a_{ij} \neq 0$ implements such nonseparability into the loss utility function (Eqn. 2.6). If both decisions are made together, the actor will still follow the unconditional ideal point. Yet, this no longer holds if the two decisions are made separately and the first decision does not match the unconditional ideal position. Then, the actor reacts by adjusting the preferences of the other topic. This adjustment resembles an actor's conditional ideal point (Enelow and Hinich, 1984, p. 18).

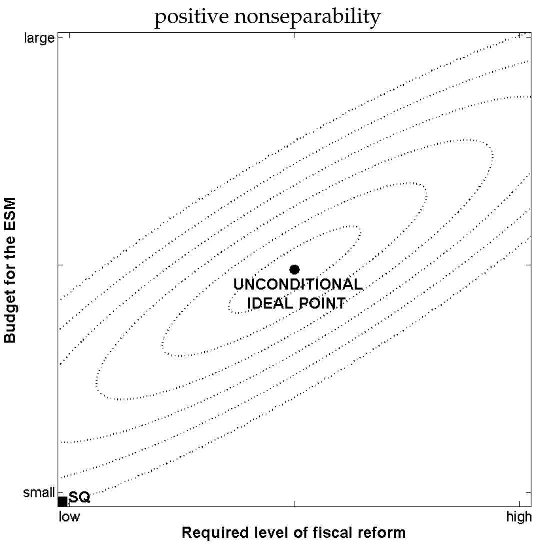
Nonseparability may be positive (rectified) or negative (directed opposite). A plain example for negative NSP is the simultaneous decision over the spending for two projects (X,Y), assuming that actors preferences are subject to either an explicit or implicit budgetary constraint. Then, they will lower their spending preference for project Y once confronted with the decision to spend more on project X than they had originally preferred (Hinich and Munger, 1997). However, the projects might just as well be complementary and, thus, interrelated by positive nonseparability. For example, consider investments into railway networks and trains. Actors would enhance their spending preference for the infrastructure once confronted with a decision to invest more into trains than they had originally hoped for. Fig. 2.3 shows the resulting indifference curves for these kinds of NSP. In this graph I use again the narrative of a Euro-group member voting on the ESM budget and level of fiscal reform.

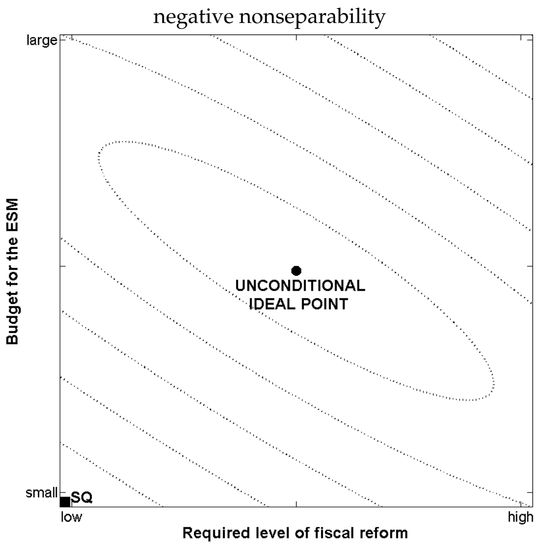
Figure 2.3: Indifference curves for nonseparable preferences

EXPLANATORY NOTE

The figures depict the unconditional ideal point and corresponding indifference curves for a member of the Euro-group who prefers a medium high budget for the ESM and medium strict fiscal requirements. The salience of issue 1 (budget for the ESM) equals 0.7 ($a_{11} = 0.7$) and the salience of issue 2 (required level of fiscal reform) equal 0.3 ($a_{22} = 0.3$). As one issue is more important, the indifference curves have the form of an ellipsis. I assume NSP by setting $a_{12} = a_{21} \neq 0$. As a result we end up with tilted or slanted indifference curves around the ideal position of the actor (cf. Hinich and Munger, 1997; Strom, 1990).

The left figure depicts positive nonseparability with $a_{12} = a_{21} = 0.8$. Here, stricter rules lead to a preference for a higher budget, and vice versa. The winset of the SQ does no cover the complete issue space. The right figure depicts negative nonseparability which is implemented by $a_{12} = a_{21} = -0.8$. Here, stricter rules lead to a preference for a lower budget, and vice versa. The indifference curves stay tilted, but resulting from the position of the SQ in the lower left corner, the winset covers the complete issue space again.





Another way of depicting NSP is using *ridge line*s. Those show "alternatives most preferred on one dimension for any given alternative in the other dimension." (Strom, 1990, p. 108) Thus, they describe the conditional ideal positions of the actors according to different outcomes of a previous decision on another issue (Denzau and Mackay, 1981). With separable preferences the preferred alternative of a dimension is not depending on other outcomes. Thus, ridge lines run only vertically or horizontally. With NSP the ridge lines are tilted. Using the two-dimensional ESM example it is easy to exemplify the calculation of the conditional ideal position. Knowing the decision made on the budget of the ESM B^* the conditional preferred level of fiscal reform $R|B^*$ is calculated as

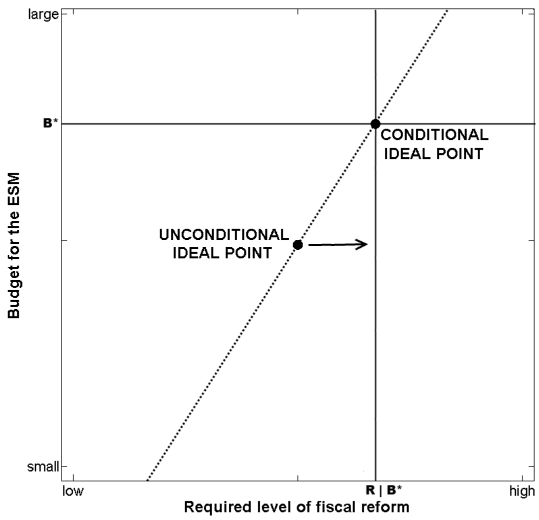
$$R|B^* = \theta_R - \left(\frac{a_{12}}{a_{22}}\right)(B^* - \theta_B) \tag{2.8}$$

Here, θ_B and θ_R describe the initial unconditional ideal position of the actor, and $\left(\frac{a_{12}}{a_{22}}\right)$ sets the amount of NSP (a_{12}) between issue 1 and issue 2 in relation to the salience of the second issue (a_{22}), i.e., the amount of fiscal reform. The computation for multidimensional space is analogous (Enelow and Hinich, 1984, p. 18). In Fig. 2.4 the ridge line for a member of the Euro-group with positive NSP are depicted.

Figure 2.4: Ridge lines for nonseparable preferences

EXPLANATORY NOTE

The figure shows the unconditional ideal point for a member of the Euro-group who prefers a medium high budget for the ESM and medium strict fiscal requirements. It also shows the actor's ridge line (dotted). As the actor has positive NSP ($a_{12} = a_{21} > 0$) a higher budget implies stricter fiscal reforms. In the figure the budget for the ESM is chosen first and fixed at B^*. This budget exceeds the initially preferred budget of the actor. Therefore, they adjust their preference with respect to the second issue (indicated by the arrow). The required level of fiscal reform is indicated as $R|B^*$. The change in the preferred level of fiscal reform leads to the members' (new) conditional ideal point.



In sum, as with the ideal position of an actor or the salience of the issues, also the degree and direction of conditionality influences the outcome of decision-making. This is represented by the different shapes of the indifference curves and the varying emergent winsets. Therefore, the utility function is supplemented by an "interaction" term. Using the culinary example to illustrate the effect this interaction only represents how well or badly the choices for wine and dish go together. This is shown in Eqn. 2.9 where d indicates the drinks and f the food served.

$$WED(\theta, x) = \sqrt{a_d\left(\theta_d - x_d\right)^2 + a_f\left(\theta_f - x_f\right)^2 + \left(a_{df} + a_{fd}\right)\left(\theta_f - x_f\right)\left(\theta_d - x_d\right)} \quad (2.9)$$

So far, two questions are still unanswered. Firstly, what upper and lower bounds exist for the magnitude of nonseparability? And secondly, does the dependency occur in both directions, so does $a_{ij} = a_{ji}$ always hold? The two following sections answer these questions.

2.4.4 The magnitude of nonseparability

This section briefly discusses the possible range of values of NSP. While the determination of a lower bound for NSP is trivial, a general definition for an upper bound is not so easy. The lower bound of 0 equals just the scenario of separable preferences where no conditionality between different issues exists. On the other hand, if such dependencies exist, how strong can they be? Like with salience, the weight attached to a dimension or to a combination of dimensions, is actor-specific. A proper definition for an upper bound should take this into account. Hinich and Munger (1997) assumed for each combination of two dimensions that the upper ceiling for nonseparability meets the criteria of

$$a_{ii} \times a_{jj} - a_{ij} \times a_{ji} > 0. \tag{2.10}$$

This assumption guarantees that the interaction effect between the dimensions $\left(a_{ij} \times a_{ji}\right)$ cannot outweigh the non-interaction part $\left(a_{ii} \times a_{jj}\right)$. In spatial terms an actor's indifference curves keep the form of an ellipsis (Hinich and Munger, 1997, p. 218). With respect to Eqn. 2.9 the criteria ensures that the term under the square root is non-negative. Obviously, assumption Eqn. 2.10 is driven by the choice of the Euclidean metric.[95] From a substantial perspective this upper bound does not define an extreme case of NSP. More precisely, there is no ex-ante reason why the interaction term should be less valuable to the actor than the two issue-specific terms.

2.4.5 The case of reciprocity

In the realm of politics the allocation of resources often depends on the correspondence of political will. For example, the transfer of decision-making authority to an autonomous government agency is often conditional upon the expected policy outcome. Yet, this does not necessarily apply in reverse. The conditionality in legislators' preferences over the characteristics of and the allocation of resources to a political program is another, although the most prominent, example. In order to model such non-reciprocity appropriately Finke and Fleig (2013) proposed a simple modification to the standard Euclidean utility function. More precisely, this implied an extension of the concept of nonseparability with respect to a) the direction and b) the reciprocity of the effect.

Many legal proposals feature a budget as well as a policy dimension. In these cases it appears straightforward to assume that legislators' spending preferences depend on the characteristics of the corresponding policies. For example, legislators' preferences for an agency's budget will be more generous if this agency pursues their political interests. Similarly, members of the IMF's Board of Governors make their decision over the size of a loan dependent on the fiscal reforms pursued by the remitting government. Whereas these statements may appear uncontroversial, political scientists tended to ignore this

[95] The empirical analyses in Chap. 3 exemplifies the practical necessity of this condition.

conditionality, because often the political decisions over the budget for, and the characteristics of, a certain policy took place simultaneously.

Commonly, the direction of NSP is assumed to be either positive or negative, as shown in Fig. 2.3. The figure discussed positive as well as negative NSP using the example of a member of the Euro-group deciding on the budget of the ESM and the required level of fiscal reform. Reciprocity assumes that nonseparability goes both ways. No matter which of the two issues is settled first, it will be followed by an adjustment in an actor's preference on the remaining issue, unless the outcome is equal to the unconditional ideal position (cf. the concept of ridge lines and Fig. 2.4).

This might not always be a useful assumption. For example, consider the case where one issue deals with the policy of a political program and another issue deals with the budget for this particular program. The spending preferences may be conditionally dependent on the policies pursued. However, the policy preferences may remain unaffected by the allocated level of spending. A political conviction stays, even if the funding provided for the favored policy changes. In terms of the ESM example this implies that the Euro-group member prefers a certain level of domestic fiscal reforms (i) which does not depend on the funding volume of the ESM; but the preferences for the size of the ESM's budget (j) most certainly depend on the fiscal reforms required for being eligible to an ESM loan. Only if the preferred policies are implemented the member is willing to vote in favor of a high budget.

Unfortunately, the idea of non-reciprocal nonseparability cannot be accommodated by simply assuming $a_{ij} = 0$ and $|a_{ji}| > 0$.[96] In the standard, reciprocal version of nonseparability (Eqn. 2.9) the sign of the interaction term $(\theta_i - x_i)\left(\theta_j - x_j\right)$ depends on both dimensions. Assuming a positive sign for $\left(a_{ij} + a_{ji}\right)$, reciprocity causes an increase in the loss utility $U_{(x_{ij})}$ if either $\theta_i > x_i$ and $\theta_j < x_j$ or vice versa. Thus, in the reciprocal case the member responds to certain policies with an increase and to certain policies with a decrease of their preferred budget.

Yet, in the case of non-reciprocal nonseparability a deviation from the preferred policies (whether too strict or too lax fiscal reform requirements) always leads to a reduced budget preference. In other words, if the member cannot get their desired policies implemented, their willingness to pay is diminished. The further the policy diverges from the members' unconditional ideal position, the smaller the most preferred budget becomes. This results from the absolute distance between an actors preferred and enacted policy in the interaction term, implemented as $|\theta_i - x_i|\left(\theta_j - x_j\right)$. The sign of the interaction term is now exclusively determined by the budget dimension j, whereas the size of the effect is determined by the policy dimension i. Simply put, divergent spending is always a bad thing, but just how bad a thing it is depends on whether or not an actor approves of the

[96]The same argument was made by an unpublished anonymous manuscript (Anonymous, 2010), which was assigned to me for reviewing, by introducing "partially nonseparable preferences" as well as "mutually nonseparable preferences".

policy being funded. Eqn. 2.11 implements this type of non-reciprocal conditionality into a standard WED.[97]

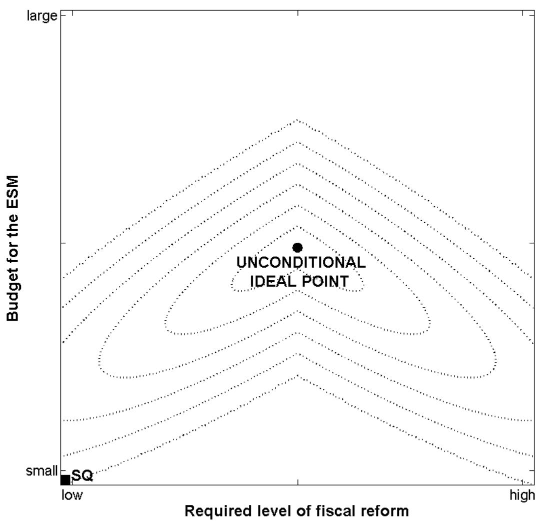
$$WED(\theta, x) = \sqrt{a_{ii}(\theta_i - x_i)^2 + (a_{ij} + a_{ji})(\theta_j - x_j)|\theta_i - x_i| + a_{jj}(\theta_j - x_j)^2} \tag{2.11}$$

Fig. 2.5 illustrates the corresponding indifference curves using the example of the Euro-group member one last time. It is easily shown that this member's tolerance for over-spending depends on the fiscal requirements, i.e., the conditions which a state has to fulfill for being eligible to an ESM loan. If the conditions are either too strict or too lax, the board members' willingness to grant more money than their unconditional spending preference is limited. In fact, if the fiscal requirements are too lax they would prefer the SQ even if the financial volume equals their unconditional spending preference, and vice versa. In other words, the combination of both issues is relevant for the decision of the board member.

Figure 2.5: Indifference curves for non-reciprocal nonseparable preferences

EXPLANATORY NOTE

The figure depicts the unconditional ideal point and corresponding indifference curves for a member of the Euro-group who prefers a medium high budget for the ESM and medium strict fiscal requirements. The salience of issue 1 (budget for the ESM) equals 0.7 ($a_{11} = 0.7$) and the salience of issue 2 (required level of fiscal reform) equals 0.3 ($a_{22} = 0.3$). NSP are implemented by setting $a_{12} = a_{21} = 0.8$. I assume non-reciprocal nonseparability, i.e., I apply Eqn. 2.11 instead of Eqn. 2.9 for calculating the indifference curves. As a result, they and the winset of the SQ are "heart"-shaped.



In politics one encounters this type of non-reciprocal NSP in different ways. Either a political actor decides over policies and corresponding budgets, or over the delegation of competences to a supranational authority and the division of power within this authority, or over the amount of regulation in combination with the extent of exceptions for affected (befriended) actors, etc. In all cases, the nonseparability is non-reciprocal in the sense that the allocation of competences, budget, etc. depends on the policy it will be used for, but not vice versa. In other words, the policy dimension defines the characteristics of the

[97] In Sec. A.2 I proof the derivation of the WED including non-reciprocal nonseparability (Eqn. 2.11).

political will, whereas the allocation dimension assigns the means with which this policy is to be enacted.

2.4.6 How to measure nonseparable preferences

The empirical estimation of NSP requires evaluating actors' utility function at more than one point of the policy space, i.e., one needs more information than just the ideal position or the first preference. This is not a trivial task when collecting data. For example, in studies using expert interviews for data collection this would increase the length of the interview. The experts would have to provide hypothetical evaluations of actors' utility function at several values. Typical questions would read as: "Assume government Z is confronted with a policy x which is unchangeable. How would this affect government Z's preference over the size of the budget y for that policy program?"

Lacy (2001b, p. 253ff) used the following arrangement of questions to identify NSP in a survey. First, respondents were asked for their unconditional first preference on an issue, e.g., "Do you think the state should increase in-come taxes, cut income taxes, or keep income taxes where they are now?" Second, respondents were asked for their unconditional first preference on another issue, e.g., "Do you think the state should spend more money to fight crime, less money to fight crime, or continue spending the same as it does now?" Finally, to identify NSP, the survey asked for the conditional response of subjects with follow-up questions: "If the state significantly cut income taxes, then would you want the state to spend more money to fight crime, less money to fight crime, or continue spending the same as it does now?"; "If the state significantly increased income taxes, then would you want the state to spend more money to fight crime, less money to fight crime, or continue spending the same as it does now?"; etc. This way of questioning had to be repeated in all combinations, and also in reversed order.

It is important to understand that this approach is a significant improvement for measuring NSP compared to Hansen (1998) who looked for trade-offs across tax and spending issues. The questions read as e.g., "Do you favor cuts in spending on national defense in order to increase spending on domestic programs like education, and highways?"; "Do you favor an increase in the federal budget deficit in order to increase spending on domestic pro-grams like Medicare, education, and highways?" These questions ask for a respondent's preference on two issues simultaneously. Thus, it is not clear if a respondent answers "no" because they do not want an increase in spending, or they do want to increase spending, but just by other means. Lacy (2001b, p. 242) therefore classified these types of questions as "double-barreled".

An innovative survey methodology was put forward by Bonica (2012). He analyzed the relationship between fiscal preferences and ideology using an interactive questionnaire. Baseline respondents were presented with the President's requested budget. Subsequently, they were asked to adjust the spending levels of the single categories with respect to their personal preferences. The "interactive" set-up incorporated trade-offs between spending dimensions. The author found that the preferences on security spending correlate with self-reported ideology. In addition, he argued that public goods can be divided into rival and non-rival government goods and services.

In summary, it is possible to measure NSP using surveys. Yet, it is time consuming, expensive and complex. The type of questions changes from concrete to hypothetical. Such theoretical questioning might very soon overstretch the nerves of any interviewee expert or layman. For that reason the correct estimation of nonseparability from real world data is a very costly and difficult undertaking. Many contributions shy away from these efforts and just make the simplifying assumption of separable preferences.

2.5 The impact of nonseparable preferences

As shown in this chapter, NSP alter theoretical expectations in three ways. Firstly, ignoring the elements in the secondary diagonal of A (cf. Eqn. 2.6) causes a misspecification of actors' utility functions. While this leaves the identification of the unconditional ideal point (or first preference) untouched, the utility assigned to the remaining alternatives will be falsely specified. As a consequence the evaluation of decision-making models may be biased. I investigate this argument in Chap. 3 by analyzing competing models of EU legislative politics and the consequences of NSP for their comparison.

> ARGUMENT 1 (misspecification): Neglected nonseparable preferences produce misspecified utility functions. This leads to distorted results of models relying on these functions.

Secondly, a strand of theoretical literature focuses on the effects of separating or combining decisions over multidimensional policies characterized by NSP (cf. Sec. 1.2). More specifically, the focus of these contributions was on the effect on the policy outcome when separating decision over nonseparable issues. Such a separation assumes different decision makers or decision-making rules for each dimension. To put it more generally, the preferences of an actor depend on exogenous factors. Well known examples are multi-level government, multistage decision-making process, etc. (Strom, 1990, p. 107ff). The purpose of this delegation or division of labor is to allow individual decision makers to invest more time, effort, expertise, etc. In theory, this should improve the quality of the results. Yet, the separation increases the complexity of the decision-making structure, thereby weakening the link between individuals' action sets and the foreseeable consequences. Also, people are subject to cognitive limitations (Stanovich and West, 2000). It is unrealistic to assume that humans are always able to take into account all restrictions and implications of a decision for other topics. Thus, the effect of NSP depends heavily on the extent and level of information an actor has (cf. Denzau and Mackay, 1981), i.e., the prevalent level of uncertainty (Bendor and Meirowitz, 2004).

Looking at committee decisions on multiple issues Ordeshook (1986, p. 252) stated that with nonseparability "the order in which the committee votes on the issues can affect the final outcome." If decisions are taken sequentially,[98] nonseparability works in favor of the individuals who decide first (Strom, 1990, p. 61ff). Anticipating the effect of their choice on the other players' preference orderings concerning the subsequent issues, they can realize higher gains when compared to simultaneous decisions (cf. Strom, 1990, p. 122ff).

> ARGUMENT 2 (first-mover advantage): If a decision which contains nonseparable preferences is taken separately and sequentially, this leads to a first-mover advantage.

Thirdly, decisions on two dimensions may be taken separately but simultaneously. In this case each of the two actors must anticipate what the other might choose. The situation gets worse when each dimension is subject to collective decision-making. One example is the case of a referendum on a set of interrelated issues (Brams et al., 1997).[99] It is important to understand that NSP are different from ordinary logrolling, the political exchange over several issues or trade across law proposal (cf. König and Junge, 2009). In contrast to a trade, the preference of an actor on one issue changes with the result of a second issue. Thus, a change in one dimension leads through the conditionality, to a re-adjustment (of preferences) in the other dimension (Strom, 1990, p. 57).

> ARGUMENT 3 (sub-optimality): If a decision which contains nonseparable preferences is taken separately and simultaneously, this leads to a sub-optimal outcome, as coordination is difficult.

Arguments 2 and 3 are analyzed in Chap. 5 and Chap. 6. Using a laboratory experiment I studied a collective group decision in comparison a situation in which the group is split into two delegations. According to the arguments, I varied the level of uncertainty as well as the voting sequence, i.e., I used sequential and simultaneous decision-making. So far, the formulations are general in nature. They address the influence of NSP as a matter of principle. When explaining the design of my experiment and depicting its specific properties, I formulate concrete hypotheses for the impact of nonseparability. Those arise from the interplay of the arguments discussed, and the experimental design. These hypotheses are directly related to statistical measures. Their relationship to the arguments is best understood as a meta-level (general argument) and sub-level (case-specific hypothesis) order.

[98]Krehbiel (1986, p. 542) argued that "legislative decision making is fundamentally sequential". The importance of sequential voting was investigated by Xia et al. (2011). Looking at multi-issue domains as an extensive-form game they proved several multiple-election paradoxes in strategic sequential voting and noted that changing the order of the issues cannot completely prevent such paradoxes.

[99]This can nicely be illustrated with the example of constructing a new town hall (cf. Sec. 2.1). A corresponding referendum would ask "Should we build it at all (yes/no), where (old town/at the river) and by which architecture (modern/classic)". At the ballot box voters preference ordering depends on the outcome of each single question which, unfortunately, they cannot know. So, people may prefer a new city hall, but only in a modern design and outside the old town.

3 The merits and costs of incorporating nonseparable preferences

To operationalize NSP is a demanding task. The additional effort can only be justified if the resulting benefit is correspondingly large. This consideration has to be made every time utility functions are specified. This chapter offers such an assessment of merits and costs. The assumption of separable preferences has been applied throughout the empirical research on EU legislative politics. The following sections now incorporate NSP into spatial models of legislative decision-making. Contrasting previous work with my results provides an answer to my first key question, whether and, if so, how much the inference of analytical research suffers when falsely assuming separable preferences. While legislative models are not the general focus of my research, they serve well as examples showing the consequences of neglecting NSP. Of course, the assessment of one empirical case is only a sample. Chap. 8 picks up again the discussion of merits and costs in a more general context.

The remainder of this chapter is structured as follows. Sec. 3.1 discusses the literature on models of legislative decision-making. Such models enable researchers to predict the outcome of legislative processes. These predictions are formalized in term of winset, core, winning coalition, etc. (cf. Sec. 2.4). Using these models it is also possible to evaluate the observed decisions of legislative bodies with respect to certain criteria. I categorize them into three groups according to their different levels of constraints: i) unconstrained bargaining models, ii) constrained bargaining models and iii) agenda-setting models. Next, I review the "Decision-making in the European Union" (further DEU) project (Thomson et al., 2006) in Sec. 3.2. The project's data set serves as basis for my empirical analysis. The effects on legislative decision-making models, when misspecifying preferences by omitting nonseparability, are explained in Sec. 3.3. Here, I include two exemplifying case studies. I start the empirical analysis in Sec. 3.4 by investigating the spread of NSP in the DEU data. Next, in Sec. 3.5 the magnitude of NSP is estimated by means of simulation techniques. Sec. 3.6 turns from the collective outcome towards the individual votes in the European Council. Finally, I summarize the finding of this chapter in Sec. 3.7.

The main empirical results of this chapter are also published in "The Merits of Adding Complexity: Conditional Preferences in Spatial Models of EU Politics" (Finke and Fleig, 2013). This article investigated the impact of including NSP into spatial models of analytical politics on comparisons of legislative models. It challenged the ambiguous assumption of separable preferences which had been applied throughout previous empirical research. The goal of this chapter is similar, thus Sec. 3.4 and Sec. 3.5 are largely associated with Finke and Fleig (2013). Compared with the publication this chapter contains additional information: the overview of legislative decision-making models has been extended (Sec. 3.1), supplementary data is discussed (Sec. 3.2), an additional case study is included (Sec. 3.3.2) and the argument for the use of modeling at the individual level is assessed (Sec. 3.6). These extensions enable a more detailed review and comprehensive analysis.

3.1 The literature on legislative models of decision-making

All legislative models of decision-making are built on a common foundation (Ferejohn and Fiorina, 1975); to evaluating the models information is necessary about the institutional environment, the preferences of the involved actors as well as their relative salience, the discussed issues, the localization of the SQ and the outcomes in policy space. Mostly, this information is collected through surveys or expert interviews. Starting from this common ground the different models emphasize various aspects in the decision-making process. This also leads to different solution and equilibria concepts.

- Dynamic exchange models: These models implement an indefinite sequence of bilateral tradings between actors. Through these (bilateral) interactions a model reaches its solution by gradual adjustments away from its initial position in reciprocal concessions. If there are no more partners for bilateral exchange found, the system is in equilibrium and the resource allocation (policy) at this point is the outcome (Coleman, 1990).[100]
- Negotiation models: These models combine multilateral negotiation with institutional constraints. To account for a variety of actors the models include a weighting variable assigning the actors different degrees of influence (Schneider et al., 2006). These models emphasize strongly the importance of institutional constraints. Typical model solution concepts are the winset, subgame-perfect equilibria or the Nash bargaining solution (NBS). Those enable the multilateral negotiation to reach a stable equilibrium (Hug, 2009). Otherwise the chaos theorem would predict unstable multiple equilibria (McKelvey, 1976; Richards, 1994).
- Procedural models: These models focus on the legislative process. Important aspects are the sequence in which different institutions are involved as well as who possesses agenda control. They aim to assess to what extend upstream decision-makers can influence the actions (and scope of actions) downstream. As typical solution concept these models use backward induction.[101]

[100] A first model of resource exchange goes back to Goode (1971).
[101] For a comprehensive overview on procedural models cf. Steunenberg and Selck (2006, p. 54ff.).

I continue by presenting the most popular implementations of these legislative models on EU decision-making.[102] These are unconstrained bargaining models (Sec. 3.1.1, implementing dynamic exchange), constrained bargaining models (Sec. 3.1.2, implementing negotiation) and agenda-setting models (Sec. 3.1.3, implementing procedural aspects).[103] All models are based on the principles of spatial modeling (cf. Sec. 2.4.2).

3.1.1 Unconstrained bargaining models

In his study of decision-making in the European Community Van den Bos (1991) presented a compromise model. This model assumes that a bargaining solution "takes all positions of member states into account, weighting these by the resources a member state can apply during the negotiation and the importance each attaches to the decision at hand" (Van den Bos, 1991, p. 176). As a consequence the model's prediction x^*_C results from Eqn. 3.1, where s_i is the relative salience actor i attaches to the issue, v_i describes actor i's relative bargaining power and θ_i is their unconditional ideal point.

$$x^*_c = \frac{\sum_{i=1}^{n} s_i v_i \theta_i}{\sum_{i=1}^{n} s_i v_i} \tag{3.1}$$

Achen (2006) proved that the prediction of the compromise model is an approximation of the asymmetric bargaining model with separable preferences if disagreement is undesirable and, therefore, not an option. This facilitates exchange and mutually adjustments. The proof is of significant importance, because it allows simulating the effect of NSP on the compromise model by applying standard spatial utility functions. The multidimensional extension of the compromise model has been termed "position exchange model" (Arregui et al., 2006). Its prediction equals the multidimensional, unconstrained and asymmetric bargaining model. The unconstrained bargaining model is the least constrained of all models. It neither considers the reference point, nor does it explicitly analyze the decision-making process.[104]

3.1.2 Constrained bargaining models

Using the DEU data set Schneider et al. (2010) calculated several versions of the symmetric and asymmetric bargaining model.[105] Briefly summarized, this bargaining model predicts that actors agree on the policy x within the winset to the SQ which maximizes

[102] For a more encompassing summary of modeling EU law-making procedures cf. Thomson et al. (2006).

[103] This excludes dynamic models such as the "expected utility model" (cf. Bueno de Mesquita, 2011; Stokman and Bueno De Mesquita, 1994) which endogenize changes in actors' negotiation position during the process. Yet, the selection points out the connection to other areas of political research. For example, "bargaining theories are omnipresent in studies on international relations" (König, 2013, p. 2) as well.

[104] An applications to the EU can be found, e.g., in Stokman and Van Oosten (1994).

[105] Baron and Ferejohn (1989) clarified the conditions under which the strategic implementation of the Stahl-Rubinstein model (Rubinstein, 1982) and the axiomatic bargaining model (Nash, 1950) lead to identical equilibrium predictions.

the product of their utility functions (further Nash product[106]). They assume the spatial Euclidean utility function presented in Eqn. 2.4. As Schneider et al. (2010) used the assumption of separable preferences Eqn. 3.2 presents the symmetric version of the model:

$$x_{nbs} = max_{x\epsilon\Theta} \prod_{i=1}^{n} (u_i(x) - u_i(SQ)) \tag{3.2}$$

The utility is defined as the difference between the distance of the actor's ideal position to the disagreement value (i.e., SQ) and the distance of the actor on the proposed policy. The constrained bargaining model is more restrictive than the unconstrained bargaining model in that it constrains the set of feasible outcomes to the winset. The latter is determined by the voting rules for each proposal, either qualified majority voting (QMV) or unanimity in the Council.

3.1.3 Agenda-setting models

Agenda-setting models produce equilibrium predictions by solving the multistage EU law-making game via backward induction. They are part of a broad rational choice literature, which regards the results of policy actions as a common result of preferences and institutions (e.g., Ostrom, 1986; Shepsle and Weingast, 1995). Hörl and Warntjen (2005) reviewed previous literature on legislative decision-making which use procedural spatial models. The authors conceded that these models have considerably enhanced the understanding of EU legislative decision-making. However, the article also explicitly addressed criticism concerning central assumptions of the approach, in particular the formation of preferences within legislative bodies.

Most prominent are the competing models by Tsebelis (1994), Moser (1997) and Steunenberg and Vught (1997). An important distinction between the models must be made with respect to the significance of European Commission and European Parliament in the legislative process under different institutional arrangements.[107] Under the consultation procedure[108] the EP may approve, reject or amend a legislative proposal proposed by the Commission. However, most of the authors agreed that the European Commission can successfully set the agenda by placing its proposal inside the winset of the SQ and the amendment proof set, i.e., the unanimity core. The powers of the EP in the co-decision procedure are theoretically more contentious.[109] This legislative arrangement devotes the same weight to EP and European Council. Here, some authors claim that the EP enjoys conditional agenda-setting powers, others argue in favor of keeping the Commission as

[106] The Nash product is the maximand of the Nash Bargaining Solution (Nash, 1950).

[107] Cf. Steunenberg (1994) for a comparison of decision-making in the European community under different institutional arrangements. He examined the consultation procedure, the cooperation procedure as well as the co-decision procedure.

[108] Cf. `http://www.europarl.europa.eu/aboutparliament/en/00c5e7159b/Consultation.html` (accessed April 28, 2013).

[109] Cf. `http://www.europarl.europa.eu/aboutparliament/en/0080a6d3d8.html` (accessed April 28, 2013).

agenda-setter, but successful amendments cannot violate the preference of the median MEP (this is called "amendment proof"). Subsequently, I follow the latter approach. Hence, the agenda-setting model not only constrains the set of feasible solutions to the winset, but furthermore includes the agenda-setting power of the Commission.

3.2 Data on decision-making in the European Union

The DEU project (Thomson et al., 2006) resembled an encompassing summary of modeling EU law-making procedures. This collaborative research project gathered information on the (ideal) position and salience of national governments, the European Commission and the EP on a total of 66 law-making proposals. To ensure a high level of salience (i.e., political relevance and contestation), the proposals were selected from the press archive Agence d'Europe[110]. The political conflict on these proposals revealed between two and six issues, resulting in a total of 192 issues discussed in the Council between January 1999 and December 2000.

The ultimate purpose of the research project was a comparative evaluation of competing models of EU law-making (cf. Achen, 2006; Schneider et al., 2006). The tremendous success of the research project manifested itself in more than 170 citations in articles and books and a still growing amount of publications using the data set (for a recent overview cf. Thomson, 2011). Motivated by this success a follow-up project has been carried out which focuses on yet another 59 legal proposals initiated after the Eastern enlargement of the European Union (henceforth DEUII, Thomson et al., 2012). The revised and expanded DEUII data includes the described information for 125 legislative proposals comprising, in total, 331 controversial issues of EU legislation between 1996 and 2008 (Thomson et al., 2012, p. 604).

Looking more into the details, the DEU data contains information for 51 multidimensional proposals which in total comprised 144 issues. They cover in equal share the Consultation and the Co-decision procedure, as well as unanimity voting and QMV[111]. The DEUII (Thomson et al., 2012) extended this collection to 101 multidimensional proposals containing 307 issues. The data shows clearly that (at least the contested) EU law-making is characterized by multidimensional (i.e., complex) decision-making.

In the DEU and DEUII data all information (i.e., individual policy position and salience of every member state for every issue) has been scaled between 0 and 100. The data includes information on the location of the SQ (enacted policy previous to the negotiations)

[110] Homepage: `http://www.agenceurope.com`.

[111] The QMV threshold refers always to the Treaty of Nice (Nice, 2001). Originally, the Treaty of Amsterdam (Amsterdam, 1997) established the European Community and defined the voting rule for the European Council of ministers (article 205 section 2). The Treaty of Nice (article 3) amended the weighting of votes in the Council and the QMV threshold. It added the requirement that "the Member States constituting the qualified majority represent at least 62 % of the total population of the Union" as well as "two-thirds of the members" (Treaty of Nice, article 3, Sec. 1.b and 1.c).

and the final outcome. All the data has been gathered by expert interviews.[112] Sec. A.3 provides a short overview on the questionnaire of these interviews. In the case of the DEU data set I also had access to the documentation of the interviews: each proposal was accompanied by a short field report, drafted by the responsible interviewer, which contains a short description of the major issues, an assessment of the member states opinions and the Commission's intention when drafting the proposal. These reports proved to be a very helpful starting point for the investigation into potential nonseparability.

The empirical analysis on the influence of NSP on legislative decision-making models was restricted to data from the DEU project (as in Finke and Fleig, 2013) for two reasons. Firstly, success and acceptance of the originally data set are undeniable (Thomson, 2011). I have no doubt that the DEUII will develop to the same prominence, but this can only be proven by time. Secondly, the investigation of NSP is difficult as the original interviews did not ask for nonseparabilities. Based on the field reports, which document the interviews in detail, this problem could be solved at least to some extent. As I had only access to the reports of the DEU data set I restricted my assessment of NSP accordingly.

Before going into the analysis of the data I take a closer look at the process of the data collection. The most important aspect of empirical data is its reliability. The in-depth expert interviews of the DEU project are recorded precisely. Also, the interviews have been conducted across all proposals with the same level of thoroughness. But can interviewing experts provide enough information to code a proposal? This includes the identification of the single issues in the proposal as well as position and salience estimates for every single member state in each issue. An alternative approach would have been to identify the single issues of each proposal through hand coding.[113] Appendix II of Thomson et al. (2006) actually contrasted the expert judgments with information found in Council documents. Benoit and Laver (2007b) set the results of expert interviews and hand-coding against each other and provided a discussion on the effects and consequences. The authors concluded that no method is generally superior to the other. When considering some caveats it is even possible to combine the two approaches.

Leuffen et al. (2012) compared different aggregation strategies for multiple sources in small-n research. The authors focused on qualitative research strategies, e.g., process tracing.[114] Overall, weighted averages of multiple origins performed best if the varying sources are independent. If independence cannot be assured the research should focus on the most precise and accurate data, depending on the trustworthiness of the source. With respect to the DEU project (nearly) all experts "had first-hand knowledge [...] and were usually participants" (Thomson et al., 2006, p. 32) in the decision-making.[115] The

[112] For each proposal up to 5 different experts were interviewed.

[113] This does not apply to the preferences of the Council members because "the written accounts of council meeting often do not detail the positions initially favored by the actors" (Thomson et al., 2006, p. 31).

[114] In general, also the mixing of qualitative and quantitative methods is possible. Here, Wolf (2010) provided an excellent discussion (including application criteria) on triangulation.

[115] Sec. 2.2 of Thomson et al. (2006) provides an overview on the criteria used in the expert selection process.

project also offers information on every single person interviewed which secured "the reliability and objectivity of measurement" (Leuffen et al., 2012, p. 3).

In general, the assumption of Euclidean space is not without problems (cf. Sec. 2.4). Yet, when taking a closer look at the proposals and issues listed in the DEU data set ,I found that they rarely represent the type of allocation-rivalry problem I describe in Sec. 2.4.2. Therefore I proceed under the assumption of Euclidean space.

3.3 The effects of misspecifying preferences

In this section I explain how the prevalent assumption of separable preferences may bias a comparative evaluation of the aforementioned legislative models. The models differ with respect to their different levels of constraints. NSP affects these constraints and leads to heterogeneous model predictions. This is because constrained models have a tendency to produce extreme predictions. For example, in terms of the DEU project, consider the most extreme case in which only one country prefers the SQ. All others prefer significant changes in the same direction. In the case of unanimity voting, this implies that constrained models will predict the SQ to prevail. Another extreme scenario illustrates the reverse effect. Assume the SQ equals zero, all member states except one are located at 50 and the one member state as well as the agenda-setter prefer full reform (100). Following the agenda-setting model, the agenda-setter chooses a proposal closest to its ideal point within the winset of the SQ and the amendment proof set i.e., the unanimity core. As a result the model predicts full-scale reform.

The crucial question therefore is to what extent a misspecification of NSP endangers the inference drawn from empirically testing spatial models of law-making. Such effects of false specification may bias the evaluation of competing models of decision-making, as the constrained models predictions hinge upon the correct specification of the winset and the unanimity core. The following two sections present two more detailed case studies.[116] Both plainly illustrate two points. Firstly, the assumption of (non)separable preferences significantly alters the winset to the SQ. Secondly, the results of the model comparison depend on the level of the assumed nonseparability.

3.3.1 Illustrative case study I

In May 1999, the Commission proposed a Council regulation for "closer dialogue with the fishing industry and groups affected by the common fisheries policy"[117] (further

[116] Both case studies discuss a proposal contained in the original DEU data (Thomson et al., 2006). The following specification of the issues is based on the corresponding field reports.

[117] The Commission Proposal COM (1999) 382 - 1999/0163/CNS is accessible at `http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=de&DosId=147501` (accessed August 29, 2012).

COM1999/163). The Council had to decide by QMV; the role of the EP had been consultative. The purpose of the proposal was to create a legal basis for supporting the representation of fishery organizations at the EU level.

Within the Council, intergovernmental conflict arose along two issues.[118] The first issue concerned the extent to which national organizations should be represented in the newly established Advisory Committee on Fisheries. Advisory Committees are widely regarded as highly influential with respect to the EU's tertiary legislation (König et al., 2010). The set of possible solutions ranged from an exclusive representation of EU-level organizations to a privileged representation of national organizations. The legal SQ provided for a mixture of national and EU-level organizations being represented. The second issue arose over the kind of legal basis which should be created for the financial support of fishery organizations to enable them to organize at the EU level. Three alternatives were on the table: i) no financial support whatsoever for fishing industry organizations, ii) the expenditure should be included in the optional "A"-part of the EU budget, iii) a legal obligation should be created to provide sufficient funding for national organizations to participate in the Advisory Committee, and this post should be included in the "B"-part of the EU budget. The SQ was optional financial support.

In the following I assumed that actors' spending preferences are conditionally dependent on the representation of interests in the committee to be created (cf. Sec. 2.4.5). Here, the argument was that representation predetermines the advisory committee's policy decisions. Therefore, any deviation from member states' ideal positions on the representation issue reduced the expected satisfaction with committee decisions and caused a decrease in their preferred level of financial support. In other words, I assumed non-reciprocal nonseparability and therefore apply Eqn. 2.11.

Fig. 3.1 depicts member states' unconditional ideal positions, the SQ ante, the winset for QMV, the actual policy outcome and the model predictions for separable preferences.[119] Member states with a powerful fishery lobby, such as Sweden, Portugal, Ireland, Spain, the UK, the Netherlands and France, preferred maximal powers for their national organizations in the Advisory Committee. While most other member states preferred the maintenance of the SQ, only Germany, the EP and the European Commission preferred to completely exclude national organizations. Moreover, the German government was alone in its pursuit off cutting the financial support to the Advisory Committee. While France, the Netherlands and the UK preferred to maintain the present level of financial

[118] In addition to the expert interviews the conflict is also well documented in Consilium (1999); this interinstitutional dossier of the European Council shows the record of the discussion of the group "Internal Fisheries Policy" on the commission proposal. It lists in detail the comments from member states (Consilium, 1999, Chap. II, p. 2ff) as, e.g., the "serious reservations [...] whether it is really necessary to adopt a new regulation and new financial measures" (Consilium, 1999, p. 3) expressed by the German delegation.

[119] The presentation of this simple two-issue case is also possible through a three-dimensional graph. This form of presentation depicts the utility aggregation in form of "utility-hills" more pictorially. However, the undulating contour blocks information. Therefore, I include the three-dimensional representation only in Sec. A.4.

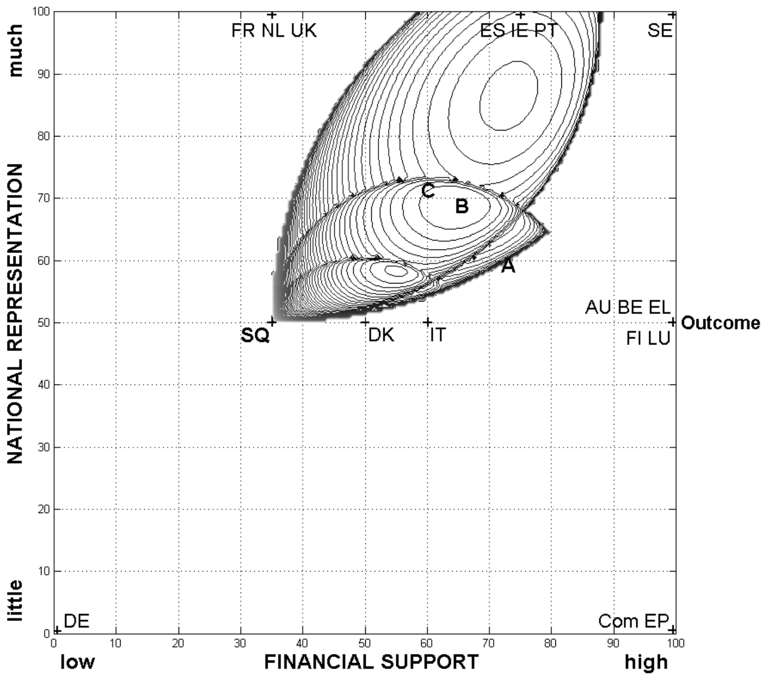
support, most other countries preferred some extension, with Austria, Belgium, Finland, Greece, Luxembourg and Sweden demanding the highest level of financial support. Not surprisingly, the European Commission (and the EP) preferred a new obligatory item in its budget.[120] The winset in Fig. 3.1 originates from the SQ, and multiple "mountains" are visible within it. These result from the different possible winning coalition. Most important are the locations of the model predictions, compared to the actual outcome. Under the assumption of separable preferences the agenda-setting model performs best.

Figure 3.1: Model predictions for COM1999/163 with separable preferences

EXPLANATORY NOTE

The figure depicts the spatial representation of member states' unconditional ideal points and model predictions for the Council regulation COM1999/163. The two issues of the proposal are the extent of national representation and the corresponding funding. The contour plots indicate the Nash product within the QMV winset of the SQ. Accordingly, the thick lines demarcate the border of the winset. Utility gains are calculated on the basis of Eqn. 2.6.

Predictions for separable preferences: A (72.4 / 59.7), B (64.9 / 69.3) and C (59.5 / 71.8).



LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom.

Subsequently, I increased the level of nonseparability. Under the assumption of Euclidean utility functions, Eqn. 2.10 provided an (empirically necessary) upper bound to the possible range of values. However, there is no ex-ante argument as to where in this range the correct value may be located. Thus, I scaled the level of nonseparability as percentage of the upper bound defined in Eqn. 2.10.

Fig. 3.2 shows the same information as before, but this time under different levels of nonseparability. Overall, increasing nonseparability shifted the winset and the prediction

[120]This does not necessarily contradict the view of the European Commission as "motor of integration" (Bailer, 2006, p. 12).

towards the de facto outcome (100/50): compared to the SQ ante (35/50) this means no change in policy but one with respect to the financial support. The shift from 0% to 50% nonseparability increased the predictive accuracy of all three models. Unchanged, it was still the agenda-setting model delivering by far the most accurate prediction. Given its unconditional ideal position and its high salience on the issues of financial support, the European Commission placed its proposal at the lower right border of the winset which, thanks to the extreme German position, is also located within the Council's amendment proof set.

Figure 3.2: Model predictions for COM1999/163 with nonseparable preferences

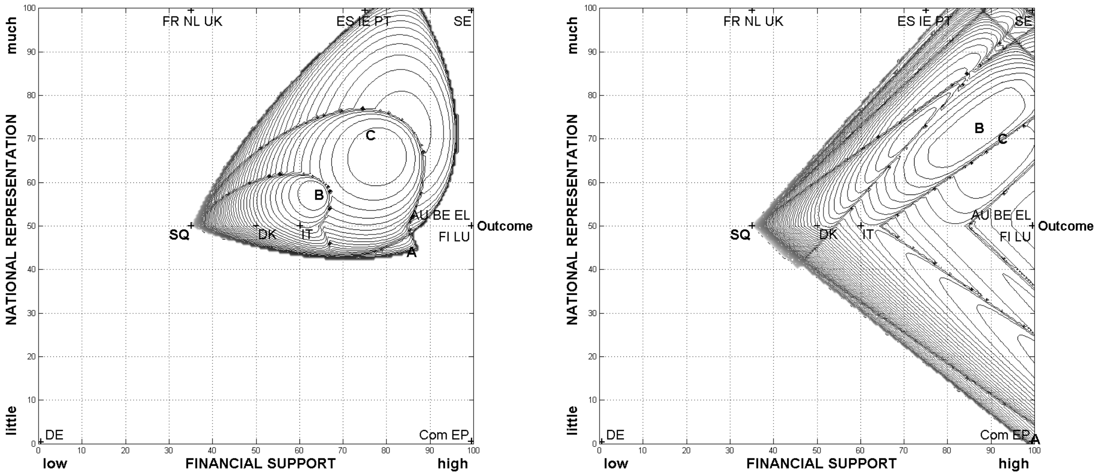
EXPLANATORY NOTE

The figures depict the spatial representation of member states' unconditional ideal points and model predictions for the Council regulation COM1999/163. The two issues of the proposal are the extent of national representation and the corresponding funding. The contour plots indicate the Nash product within the QMV winset to the SQ. Accordingly, the thick lines demarcate the border of the winset. Utility gains are calculated on the basis of Eqn. 2.11.

Predictions for NSP, when the amount of NSP is determined by Eqn. 2.10 and amounts to ...

... 50% of the maximum, i.e., $a_{12} + a_{21} = \frac{a_{11}+a_{22}}{2}$.
A (85.2 / 44.5); B (63.9 / 57.5); C (75.7 / 71.4).

... 100% of the maximum, i.e., $a_{12} + a_{21} = a_{11} + a_{22}$.
A (99.9 / 59.7); B (86.7 / 73.1); C (92.0 / 70.5).



LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom.

Under full-blown nonseparability, the agenda-setting model's prediction proposal was located very close to the Commission's ideal point. At this level of nonseparability, the constrained bargaining model proved to be most accurate. The example clarifies that a different degree of nonseparability leads to different model predictions and different conclusions when assessing model performances.

3.3.2 Illustrative case study II

In November 1999, the European Commission proposed a regulation amending the "common organisation of the market in bananas"[121] (further COM1999/582). The purpose of

[121] The Commission Proposal COM (1999) 582 - 1999/0235/CNS is accessible at `http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=en&DosId=153098` (accessed December 20, 2012).

the proposal was to modify the existing quota import regime for bananas, because successive rulings from the World Trade Organization (WTO) found certain aspects of the previous regime[122] not to be in conformity with WTO standards (EC, 2000). As bananas are a common commodity the proposal was important for many countries that export bananas to the EU as well. Thus, in addition to the EP Ecuador and the U.S. participated consultatively in the debate.[123] I again chose a Council decision reached by QMV, because the minor blocking-possibilities of single states enable me to highlight the effect of NSP clearer.

The Council discussed two controversial issues of this regulation in particular. The first issue was the type of import regime that would be adopted. The set of discussed frameworks ranged from i) keeping the SQ (which would violate the WTO ruling) ii) over a simple increase of the current quotas iii) to a transformation of the system into a liberal regime based on a flat tariff. In this continuum the SQ represents a protectionist quota system. The second issue arose over a transitional period during which a tariff quota would apply. The objective of the transitional regime was to enable banana-producing regions to make appropriate adjustments to a freer market. While the general idea was not contested its duration was. The different options asked for a transitional system until i) the year 2000, ii) the year 2006, iii) the year 2010 or an unlimited system. The SQ ante referred to unlimited usage as the date for a new import regime had to be set by the proposal.

With respect to this proposal I assumed that actors' preferences are conditionally dependent. This judgment was reinforced by the corresponding field report of the DEU project and the expert statement ascertaining that "the type of system and the timing are intimately interrelated because the agreement on one may accept the agreement (or disagreement) on the other one and vice-versa" (Thomson et al., 2002, p. 7). More specifically, the degree of liberalization determined how much transitional time was considered necessary to adapt. Furthermore, to set a date for the new import system influenced how much liberalization was deemed appropriate. In other words, I assumed mutually positive and reciprocal nonseparability and apply Eqn. 2.9.

As in the previous illustrative case study I start by depicting in Fig. 3.3 member states' unconditional ideal positions, the SQ ante, the winset under QMV, the actual policy outcome and the model predictions for separable preferences.[124] The graph shows that, again, member states lined up according to interests of powerful national lobbies. Eu-

[122] The previous banana import system was based on the Council Regulation (EEC) No 404/93, as amended in 1998 by Regulation (EC) No 1637/98.

[123] In addition to the expert interview the proposal is also well documented in Consilium (2000); this interinstitutional dossier of the European Council explains in detail the proposal intentions, the amended paragraphs and the "numerous close contacts with supplier countries and other interested parties" (Consilium, 2000, p. 2).

[124] I also include for this case study a three-dimensional representation in Sec. A.4. This presentation depicts the utility aggregation in form of "utility-hills" more pictorially but the undulating contour blocks information.

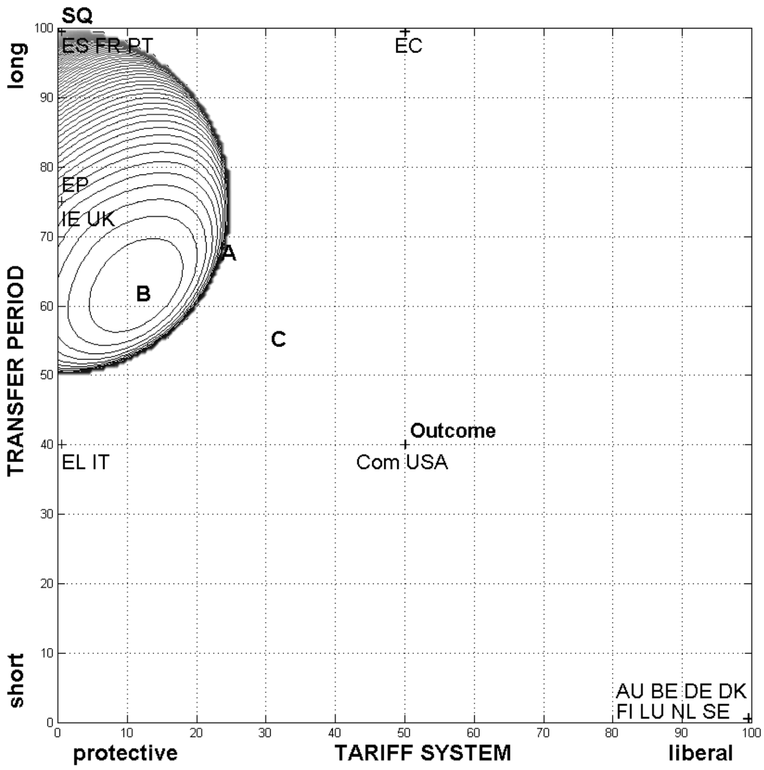
ropean producers of bananas, i.e., France, Spain and Portugal, preferred the current SQ (0/100). It was similar in the case of Ireland and the UK. Both host large companies working in this sector.[125] Italy also favored the SQ despite not producing any bananas. The reason is that Italy preferred subsidies for bananas to diminish the corresponding (fixed) EU budget to protect other vegetables of importance for Italy.

Figure 3.3: Model predictions for COM1999/582 with separable preferences

EXPLANATORY NOTE

The figure depicts the spatial representation of member states' unconditional ideal points and model predictions for the Council regulation COM1999/582. The two issues of the proposal are the import system to be adopted for the commerce and distribution of bananas within the EU countries and the year in which the new system will take effect. The contour plots indicate the Nash product within the QMV winset to the SQ. Accordingly, the thick lines demarcate the border of the winset. Utility gains are calculated on the basis of Eqn. 2.6.

Predictions for separable preferences: A (24.1 / 68.3), B (11.8 / 62.4) and C (31.3 / 55.8).



LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EC = Ecuador; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom; USA = United States of America.

A rather large coalition of countries (with no economic relations to banana import) was located at the other extreme, promoting a liberal regime in favor of cheap consumer prices. The report also outlined the intentions of Ecuador and the U.S. which, not surprisingly, preferred a liberalization of the current system. The commission was located at a medium position on both issues, perhaps having positioned itself as a mediator between the extreme coalitions. Only one winset originated from the SQ. The contrast between the

[125]In the corresponding field report the interviewed expert clarified that in both countries important companies trade directly with ACP countries (African, Caribbean and Pacific Group of States).

two groups made it hard to reach an agreement at all. Under the assumption of separable preferences the constrained bargaining model performed best.

In Fig. 3.4 I scaled the level of nonseparability as percentage of the upper bound defined in Eqn. 2.10. The figure provides the same information as before under different levels of nonseparability. Increasing the amount of NSP shifted the winset and model prediction towards the de facto outcome (50/40) which means a moderately strong liberalization within a medium transfer period. Interestingly, the outcome exactly matched the position of the U.S. and the European Commission.[126]

Figure 3.4: Model predictions for COM1999/582 with nonseparable preferences

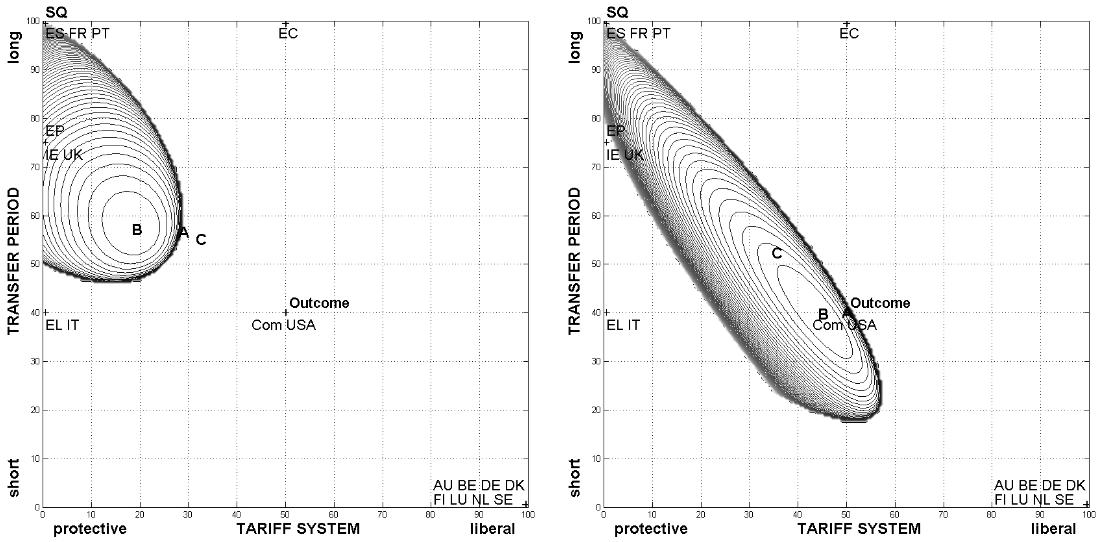
EXPLANATORY NOTE

The figures depict the spatial representation of member states' unconditional ideal points and model predictions for the Council regulation COM1999/582. The two issues of the proposal are the import system to be adopted for the commerce and distribution of bananas within the EU countries and the year in which the new system will take effect. The contour plots indicate the Nash product within the QMV winset to the SQ. Accordingly, the thick lines demarcate the border of the winset. Utility gains are calculated on the basis of Eqn. 2.9.

Predictions for NSP, when the amount of NSP is determined by Eqn. 2.10 and amounts to ...

... 50% of the maximum, i.e., $a_{12} + a_{21} = \frac{a_{11}+a_{22}}{2}$.
A (28.4 / 57.1); B (18.8 / 57.7); C (32.0 / 55.7).

... 100% of the maximum, i.e., $a_{12} + a_{21} = a_{11} + a_{22}$.
A (50.0 / 40.0); B (44.6 / 39.4); C (35.1 / 52.8).



LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EC = Ecuador; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom; USA = United States of America.

The shift from 0% to 50% nonseparability increased the predictive accuracy of the agenda-setting and the unconstrained bargaining model. But still the constrained bargaining model provided the most accurate prediction. The increase in nonseparability clearly reshaped and enlarged the winset of the QMV decision. This corresponds to the discussed changes of indifference curves (Sec. 2.4.2) which become more and more tilted as NSP increase. Under full nonseparability the winset extended further and was adjacent to the

[126]The field report clarified that the depicted outcome represents the in 2006 valid agreement. Later on, the system was further liberalized. France, Spain and Portugal gave up their resistance and a flat tariff regime was introduced.

de facto outcome. Thus, the agenda-setting model predicted the outcome exactly. The performance of the unconstrained bargaining model also further improved. Its prediction was now very close to the outcome, in contrast to the constrained bargaining model which now performed worst. This example further illustrates the effect of NSP on the different legislative models.

3.3.3 Hypotheses

The aim of the DEU research project was a performance comparison of different legislative models. Such a comparison helps to determine the importance of processes and institutions in relation to bargaining power and dynamic interactions. In the existing literature the predictive accuracy of competing models is evaluated by comparing the mean average error (MAE) at the issue or at the proposal level (e.g., Schneider et al., 2010; Thomson, 2011).[127] As an alternative, and to account for case-specific idiosyncrasy, predictive accuracy was also evaluated using pairwise comparison and so-called hit rates, i.e., predictions within a predefined, narrow margin to the real outcome (cf. Achen, 2006).

The DEU project claimed two findings (Schneider et al., 2006, p. 303): firstly, unconstrained models (such as the unconstrained bargaining model) reveal a smaller MAE. Secondly, the constrained models (e.g., the agenda-setting model) reveal a better hit rate. Yet, as the DEU project did not look into possible conditionalities in their multidimensional law proposals, the final conclusions of their research may be biased. The sensitivity to the correct specification of winset and core depends on correctly specified utility functions. This brings me to the following hypotheses.

> HYPOTHESIS 1: Falsely assuming separable preferences decreases the predictive power of models and constrains the set of feasible outcomes.
>
> HYPOTHESIS 2: Agenda-setting models are particularly vulnerable because their predictions are placed at the boundaries of the falsely specified set of feasible outcomes.

A ubiquitous assumption of separable preferences can lead to biased predictions and wrong implications about the logic of policy-making. This directly follows from ARGUMENT 1 formulated in Sec. 2.5: “Neglected nonseparable preferences produce misspecified utility functions. This leads to distorted results of models relying on these functions.” I argue that the conditionality of preferences should be taken into account. This holds for the collection of such data as well as for the empirical analysis.

My analysis proceeded in three steps. Step one identified the occurrence and direction of NSP in the DEU data set. Working on from there, step two used the coded NSP scheme as independent variable for the model comparison. Thus, revealing the different impact

[127] The MAE is calculated as the average distance between model prediction and the facto outcome across all observations.

of nonseparability on the models' predictive accuracy. As the interviewers did not ask for NSP it was not possible to encode distinct values for NSP.[128] Thus, in step three I simulated model predictions for different levels of NSP.

3.4 The extent of nonseparable preferences in EU law-making

In this section I investigate the extent of NSP in the DEU data set. Therefore, I filled in the secondary diagonal of the matrix A (Eqn. 2.6) with regard to the existence and the direction of NSP. Furthermore, I identified whether the effect is reciprocal or not. If so, the nonseparability term had to be modified by using the absolute distance of the allocation dimension as introduced in Eqn. 2.11.

The coding itself was conducted by three graduate students independently of one another. All students received a thorough introduction into the issue of nonseparability. The coding focused on three aspects: Firstly, it was identified whether or not for actors' preferences over two issues nonseparability exists at all. Secondly, the direction of nonseparability (positive or negative) was added. Third, it was determined whether the NSP are reciprocal or not, i.e., whether the issues can be classified to resemble a policy and an allocation dimension. Inter-coder reliability was approximately 90% with respect to the existence and direction of nonseparability and 75% with respect to the reciprocity.[129]

Tab. 3.1 summarizes the results of the coding efforts. NSP were discovered in half of the multidimensional proposals and in slightly more than half of the issues.

Table 3.1: Extent of nonseparability in EU law-making

EXPLANATORY NOTE
The table lists the extent of nonseparability in the DEU data set (Thomson et al., 2006). The observations are structured according to the number of issues within a proposal.

No. of dimensions	Total no. of proposals	Total no. of issues	No. of proposals affected by NSP	No. of issues affected by NSP
2	25	50	14	28
3	15	45	7	18
4	7	28	4	15
5	3	15	2	10
6	1	6	1	6
Total	51	144	28	77

These values do not imply, however, that half of the cells in the secondary diagonals of matrix A were affected. For example, the proposal for a regulation of "Audiovisual industry: development, distribution and promotion of works" (CNS/99276) aimed at subsidizing certain enterprises in the audiovisual industry. Issue 1 was concerned with

[128] However, the field reports contain information about the interrelation of issues. This information was used to construct the NSP scheme in step 1.

[129] As stated before, the main empirical results of this chapter are published in Finke and Fleig (2013). The coding scheme resembles the empirical basis of this article. Accordingly, the correct coding of disputed issues in the NSP scheme has been discussed and settled by the authors.

the amount of money allocated to the project, whereas issues 2, 3, 4 and 5 dealt with the distribution of the money. Hence, NSP were suspected to exist between issue 1 and each of the other four issues. Preferences between issues 2, 3, 4, and 5 were separable. Nevertheless, Tab. 3.1 reports all fives issues as being somehow affected by nonseparability.

Sec. A.5 provides a complete list of the NSP coding, including a short description of proposals and issues. Admittedly, it is unfortunate that it is only possible to theorize on and not prove the existence of NSP. However, in much the same way all previous empirical studies using the DEU data could only assume but not prove separable preferences.

Next, the NSP scheme served as independent variable for the model comparison. The dummy variable "NSP" equaled 1 if an issue had been affected by NSP, otherwise it equaled 0. Tab. 3.2 compares the predictive accuracy of the three models introduced above (Sec. 3.1). On average, it repeats the main findings of the DEU project. Models highly constrained by procedural assumption reveal a higher MAE than the unconstrained bargaining model. Following Tab. 3.2, this difference in the models predictive accuracy does not depend on the existence of NSP. As the table contrasts findings on issue and proposal level, I reveal the number of observations N in each case separately.

Table 3.2: Model comparison of mean average error on issue and proposal level

EXPLANATORY NOTE
The table lists the MAE on issue and proposal level for the different models. It separates predictions with and without NSP. The standard deviation (SD) is shown in parentheses.

LEVEL OF ANALYSIS	MEAN AVERAGE ERROR without NSP Issue N=63	MEAN AVERAGE ERROR without NSP Proposal N=23	MEAN AVERAGE ERROR with NSP Issue N=81	MEAN AVERAGE ERROR with NSP Proposal N=28
Unconstrained bargaining model	24.47 (18.73)	29.67 (19.01)	24.57 (26.69)	28.67 (17.73
Constrained bargaining model	30.66 (25.44)	28.99 (24.89)	28.22 (24.35)	31.59 (25.44)
Agenda-setting model	29.59 (30.01)	29.60 (30.88)	27.48 (26.77)	32.66 (32.48)

Reporting standard errors and confidence intervals (CI) has been criticized because the policy positions in the DEU data have been measured using different scales (e.g., Achen, 2006). Hence, standard errors should be interpreted with caution. Moreover, Junge (2010) argued that the error term should be modeled at the level of individuals' utility function because theoretical models themselves are based on the actions of individuals. Although the merits of this approach seem clear, so far, it is not the accepted new modeling standard.[130] Sec. 3.6 picks up this argument again and provide an overview and outlook on voting models based on individual data.

The relative accuracy changed when comparing the three models pairwise. The regression results in Tab. 3.3 reveal that the bias resulting from NSP is significantly advantageous to the unconstrained bargaining model. The regression models deployed the stan-

[130] This assessment of the literature excludes those studies which use the DEU data to explain choices of individual actors (e.g., Cross, 2012; König and Junge, 2009).

dard set of control variables used for model comparison in the DEU project (cf. Schneider et al., 2010). This included two variables controlling for a potentially systematic measurement error, namely scale and skewness of the preference distribution. The most powerful predictor to explain the models' different errors was a dummy for whether or not the SQ has in fact been changed. As argued in Sec. 3.3.3, constrained models are very sensitive to the veto power of individual member states. Accordingly, it comes as no surprise that the unconstrained bargaining model was significantly more powerful in predicting policy change. Finally, the dummy for NSP turned out to be significant for a comparison between the unconstrained bargaining model and the agenda-setting model, as well as for a comparison between the unconstrained and constrained bargaining model.

Table 3.3: Nonseparability effect on relative model performance

EXPLANATORY NOTE
The table shows the effect of nonseparability on relative model performance at the issue level. Statistically significant (two-tailed) at the 0.1 level *, at the 0.05 level ** and at the 0.01 level ***.

N = 144	RELATIVE MODEL PERFORMANCE					
	△Error: unconstrained bargaining model - constrained bargaining model		△Error: unconstrained bargaining model - agenda-setting model		△Error: constrained bargaining model - agenda-setting model	
	Coef.	SE	Coef.	SE	Coef.	SE
Scale (1 if dichotomous)	-2.63	4.96	-11.77	7.68	-9.98	6.62
Number of Issues	-1.43	1.49	0.11	2.31	1.32	1.96
Skewness of preference distribution	-3.24*	1.66	1.88	2.57	5.12**	2.19
Voting rule (1 if QMV)	6.54*	3.73	8.17	5.78	1.62	4.91
Procedure (1 if Co-decision)	2.60	3.48	-1.18	5.19	-3.73	4.57
Reform (1 if SQ prevails)	18.88***	8.33	23.94***	6.70	5.07	5.69
NSP	-6.76**	3.41	-9.35*	5.27	-2.59	4.48
Constant	-3.35	5.94	-10.87	18.67	-7.52	7.81
Adjusted R^2	0.16		0.10		0.02	

3.5 The magnitude of nonseparability

So far, the evidence shows that nonseparability occurs in EU law-making (Tab. 3.1). It also biases the comparison of constrained to unconstrained models if misspecified (Tab. 3.3). The ultimate question is: is it possible to accommodate NSP to ensure a fair model comparison? In the most ideal case, the information for each cell in the secondary diagonal of matrix A could be gathered via expert interviews, just as the values for salience in the main diagonal. Of course this would increase the length of the interview, as the experts would have to provide hypothetical evaluations of actors' utility function at several values (as discussed in Sec. 2.4.3). But it is simply impossible to gather actor-specific data on the strength and direction of any potential nonseparability in retrospect. And the interviewers of the DEU project did not ask specifically for NSP when conducting their expert interviews. This section will therefore discuss the simulation of model predictions for different degrees of NSP.

Above, the elements in a secondary diagonal have been coded with respect to three criteria (cf. Sec. A.5): the existence of nonseparability, its direction (positive or negative) and whether the nonseparability is reciprocal or not. All this information was used to enhance the simulation of different values of NSP. The strength of nonseparability was simulated for values between 0 and the upper bound defined in Eqn. 2.10, i.e., $a_{ii} \times a_{jj} - a_{ij} \times a_{ji} > 0$. This upper bound varied according to the relative salience each actor attached to the two issues in question. To ensure the comparability of the results the simulated strength of nonseparability was operationalized as a percentage of these maximal values.

This approach has several limitations. The simulations did not account for constellations in which different actors' preferences were characterized by either different levels or different directions of nonseparability. Due to the discontinuity of the object function it was almost impossible to employ typical optimization methods for constrained nonlinear optimization problems such as Sequential Quadratic Programming (SQP) in order to identify the global optimum.[131] A proper choice was the utilization of a derivative-free global optimization algorithm. I implemented this algorithm in MATLAB (Version 7.12.0.635) based on the idea of Perttunen et al. (1993). Using MATLAB's "Global Optimization Toolbox" (Version 3.2-R2011b)[132] and its related extension "Genetic Algorithm and Direct Search Toolbox"[133].

Fig. 3.5[134] shows each model's MAE evaluated at different levels of nonseparability. Overall, 40 of the 81 issues revealed NSP in a non-reciprocal relation. Accordingly, the figure depicts the results for two sub-samples. Those 41 issues categorized with reciprocal NSP reveal an improved predictive accuracy for levels of nonseparability up to approximately 65% of the maximum. This improvement was weakest for the unconstrained bargaining model, whose predictive success was outperformed by the constrained bargaining model for levels of nonseparability larger than approximately 45% of the maximum. For this sub-sample the predictive accuracy of the agenda-setting model equaled the unconstrained bargaining model for levels of nonseparability larger than 50%. For the second sub-sample of 40 issues which were affected by non-reciprocal NSP the results differ. Assuming separable preferences all three models clearly performed worse when compared to the other sub-sample. This finding suggests that modeling a non-reciprocal relation between issues is an error-prone undertaking. Assuming higher levels of nonseparability left the predictive accuracy of the bargaining as well as the unconstrained bargaining model unchanged. Yet, the predictive accuracy of the agenda-setting model increased

[131] Cf. Kröning and Strichman (2008) for a broader overview on procedures for automated verification and reasoning, theorem-proving, compiler optimization and operations research.

[132] Source: `http://www.mathworks.de/products/global-optimization` (accessed April 4, 2012).

[133] The corresponding user's guide (version 1) can be obtained from `http://www.mathworks.com/help/releases/R13sp2/pdf_doc/gads/gads_tb.pdf` (accessed April 4, 2012).

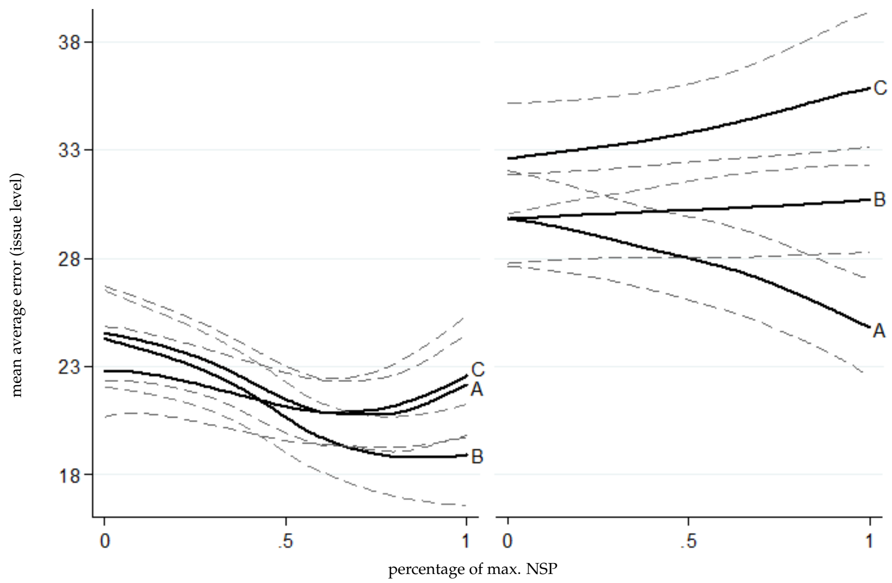
[134] The DEU data applies different measurement scales. Therefore CI, as used in the figures in this section, should be interpreted with caution (Achen, 2006).

drastically. This increase became significant for levels of NSP larger than 90%, a level at which the agenda-setting model clearly outperformed its two competitors. Assuming higher levels of NSP the agenda-setting model successfully overcame the SQ bias.

Figure 3.5: Mean average error per issue at different levels of nonseparability

EXPLANATORY NOTE

The figure depicts each model's MAE per issue at different levels of nonseparability and includes 90% CI. The graphs are the result of a local polynomial smoother using an Epanechnikov kernel (Epanechnikov, 1969) with a degree of 0. It is best interpreted as a moving average through the simulated data. The 41 issues categorized with reciprocal NSP are depicted in the right and the 40 issues categorized with non-reciprocal NSP in the left figure.



LABELS

A = agenda-setting model; B = unconstrained bargaining model; C = constrained bargaining model.

Fig. 3.6 compares two models predictive accuracy by subtracting their predictive errors for each individual issue (paired comparison). The main results resemble the findings of Tab. 3.3 and Fig. 3.5.

In Sec. A.6 in Fig. A.3 I depict also each model's MAE and the paired comparison at the proposal level. The errors are standardized on the maximal error size, which depends on each proposal's dimensionality.[135] These figures were placed in the appendix because the main results resemble just the findings at the issue level.[136] For low levels of nonseparability the unconstrained bargaining model outperformed its competitors. With respect to the minority of 7 proposals which remained unaffected by any allocation dimension, an increase in nonseparability up to approximately 80% caused an increase in all models' predictive accuracy. With respect to the majority of 21 proposals affected by an allocation dimension, the constrained bargaining model and the agenda-setting model gained in predictive accuracy. More specifically, the agenda-setting model clearly outperformed its competitors for higher levels of nonseparability.

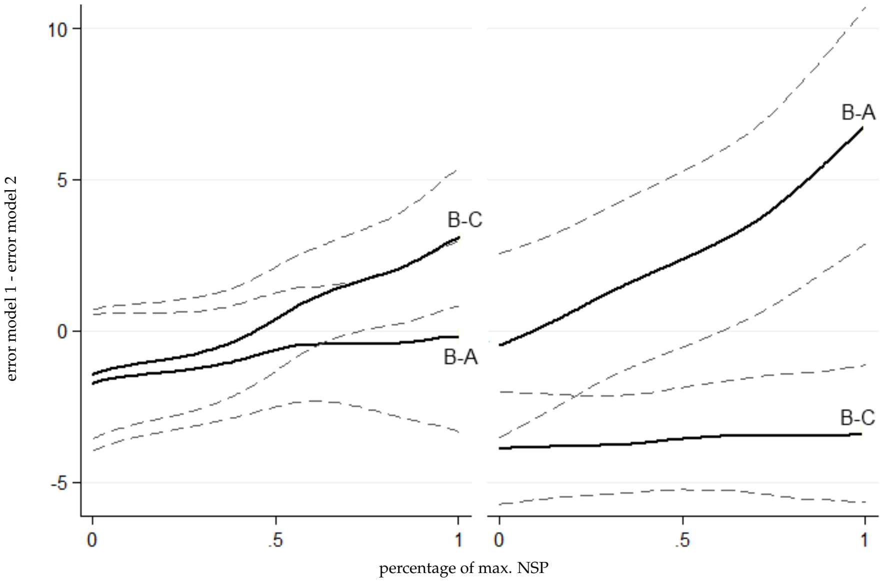
[135] The maximal error size increased by the factor of 100 for each additional dimension. This is due to the scaling of all individual issues between 0 and 100 in the DEU and DEUII data.

[136] In addition, the different spatial dimensionality prevents the application of the hit rate as an criterion for the predictive power of the models at the proposal level.

Figure 3.6: Comparison of model's predictive accuracy at the issue level

EXPLANATORY NOTE

The figure depicts models relative predictive accuracy. The graphs are the result of a local polynomial smoother using an Epanechnikov kernel (Epanechnikov, 1969) with a degree of 0. It includes 90% CI and is best interpreted as a moving average through the simulated data. The 41 issues categorized with reciprocal NSP are depicted in the right and the 40 issues categorized with non-reciprocal NSP in the left figure.



LABELS

A = agenda-setting model; B = unconstrained bargaining model; C = constrained bargaining model.

An alternative method used to evaluate the predictive accuracy of competing decision-making models is the calculation of hit rates (Achen, 2006). The underlying idea is that models should be able to provide forecasts with a reasonable degree of precision. Hence, the quantity of interest is the percentage of cases in which a model predicts with an accuracy of less than 1%, 5% or 10% of the maximum divergence. Given the data at hand the corresponding thresholds were 1 point, 5 points and 10 points on the 100 point scale.

Tab. 3.4 presents the corresponding numbers for the three models. Comparing the second to the third column it is obvious that those cases for which the separability assumption was suspected to be false, a lower hit rate at all three levels of precision is revealed. Please keep in mind that the agenda-setting model predicts perfectly when the final position equals the Commission's position which, however, must be in the amendment proof set. As a consequence, and starting from a very low number of correct predictions, the two bargaining models showed a stronger increase in highly accurate hit rates when non-separability was tuned in. This relative advantage vanished once the precision threshold was altered from 1 per cent to 10 per cent. The test for differences in binomial proportions showed that only the agenda-setting model revealed a statistically significant improvement in the hit rates under all three thresholds.[137] By contrast, the unconstrained and

[137]The corresponding test statistic is calculated as $z = \frac{p_{100} - p_0}{\sqrt{\frac{p_0 \times (1-p_0)}{n}}}$ with n as the number of observations, p_0 as the hit rate under NSP = 0% and p_{100} as the hit rate under NSP =100% (cf. Sprinthall, 2011, Chap. 4 and 5). The null-hypothesis assumes no differences in proportions and is rejected at an 0.1 level for $z \geq 1.28$.

constrained bargaining model revealed even lower hit rates at the 10 per cent threshold once nonseparability had been accounted for.

Table 3.4: Model hit rates

EXPLANATORY NOTE
The table lists models' hit rates at different levels on nonseparability. The z-Test is used to determine significant differences in performance.

	MODEL HIT RATES				
	Cases without NSP	NSP = 0%	NSP = 50%	NSP =100%	z-Test comp. NSP 0 - 100%
DIVERGENCE <1%					
Unconstrained bargaining model	13.1	4.3	7.9	8.4	6.82
Constrained bargaining model	16.9	7.9	11.1	15	6.52
Agenda-setting model	20	13.7	15.1	16.9	2.10
DIVERGENCE <5%					
Unconstrained bargaining model	21.5	18.50	18.40	18.20	0.16
Constrained bargaining model	24.3	18.10	19.70	22.30	2.09
Agenda-setting model	26.9	19.10	20.90	23.80	2.20
DIVERGENCE <10%					
Unconstrained bargaining model	40.8	38.7	36.4	33.3	1.50
Constrained bargaining model	39.7	36.1	35.3	35.1	0.28
Agenda-setting model	33.1	27.4	29.4	33.4	1.98

A comparison of the unconstrained bargaining model and the agenda-setting model showed that under the false assumption of separable preferences the former outperformed the latter by 38.7 to 27.4 per cent. However, under the assumption of a very high level of nonseparability the agenda-setting Model delivered 33.4 per cent correct predictions, whereas the performance of the unconstrained bargaining model dropped to 33.3 per cent. These results further supported the argument according to which falsely-assumed separability gives a systematical advantage to unconstrained models.

In sum, the findings are threefold. Firstly, for those issues unaffected by reciprocal NSP an increase in nonseparability up to approximately 70% improve the predictive accuracy of all three models. Secondly, this improvement is strongest for the constrained and weakest for the unconstrained bargaining model. Thirdly, a drastic increase in the predictive accuracy of the agenda-setting model for issues affected by non-reciprocal NSP is observed. The agenda-setting model emerges as the most powerful model for levels of nonseparability higher than 50%. Overall, in line with Finke and Fleig (2013) it is safe to conclude that the advantage of unconstrained models reported in the literature (cf. Sec. 3.1) has been (at least) reinforced by falsely assuming separable preferences. The simulation of NSP diminishes the advantage, enables a fairer model comparison and underscores the agenda-setting power of the European Commission.

3.6 The nonseparability in individual votes

This section takes up again the argument that models which are based on individual decisions should also model the error term of utility functions at that level (cf. Junge, 2010). The modeled source of error affects predictions and performance of a model (cf. Signorino, 1999). According to Signorino (2003), model error of strategically acting individuals can result from various types of uncertainty as bounded rationality of agents, actor's private information and regressor error or strategic error in specification (Signorino and Yilmaz, 2003).

This argument holds true for all theoretical models of decision-making and implies that model predictions should not only be applied at the aggregate but also the individual level. Thus, the model would not only predict the collective outcome but also deliver a detailed record of each single actor voting in favor or against a proposal. The accuracy of the single votes could then be assessed with respect to the implemented degree of nonseparability. In the ideal case this would be done actor-specific. Unfortunately, this step was not feasible for my research as the DEU data did not contain enough observations for an individual assessment. In the remainder of this section I discuss in detail the problems I encountered. I also show what previous contributions have accomplished in this area and detail some options for future research.

The problems with regard to an individualized analysis were twofold. Firstly, the expert interviews of the DEU project used to assemble the NSP scheme contained, if any, only general remarks on conditionality. To supplement this data by individual co-variates as sector-specific net contributions to the EU budget only resulted in crude approximations. Thus, the coding could only operationalize NSP for all actors the same way. Secondly, the total number of usable observations in the DEU data set was too small. For the examination of individual voting behavior only majority decisions can be used as in unanimous decisions a yes or no distinction cannot be made. In addition, these decisions must be reached with certain level of disagreement. A majority decision taken unanimously contains not more information as a unanimous decision in the first place.

Looking more into details, the small number of observations is not caused by lack of accessibility as all voting records of the EU Council can be publicly obtained. Pursuant to article 207 (3) of the Treaty of Amsterdam (Amsterdam, 1997) the results and explanations of the EU Council legislative votes are made public. They are published as Council minutes[138] which summarize the discussion of the Council meetings and contain the voting records as addenda (for a more detailed description cf. Hagemann and De Clerck-Sachsse, 2007). In addition, the General Secretariat provides a monthly summary of Council acts (legislative as well as non-legislative), which includes the results

[138] All Council minutes since 1999 can be accessed at `http://consilium.europa.eu/documents/legislative-transparency/council-minutes`.

of all ballots, explanations of voting and general statements.[139] Thus, secrecy is not the reason for the insufficient amount of data. In fact, many previous studies used data from the voting record of the EU Council of ministers. These studies analyzed the impact of related EU institutions, different decision rules and party affiliation on Council decisions (for an overview cf. Hagemann, 2007, p. 280-281). The contributions used various theoretical (e.g., spatial) models to approach the data. Hagemann (2007) addressed the measurement of policy preferences in the Council of Ministers and which research is suited to its voting data. The study focused on the type of data that is in fact available from the Council. Assessing different analyzing methods a simulation-based Bayesian Monte Carlo Markov Chain (MCMC) model proved to be most adequate. Thus, also the appropriate theoretical decision model has already been determined.

However, most decisions of the EU Council are reached in a "culture of consensus" (Hagemann and De Clerck-Sachsse, 2007, p. 3). This results in predominantly unanimous decisions and a level of disagreement not sufficient for individual voting assessment. König and Junge (2009) studied this unusual prevalence of consensus based on voting preferences and logrolling opportunities of the member states on 48 Commission proposals. The authors found that even member states with veto power support Commission initiatives they prefer less than the SQ. Such solutions are made possible by logrolling across domain-specific or simultaneously decided proposals. They are also driven by the Commission's attempts to avoid a divided Council. Another facilitator of the consensus decisions is identified by Mattila (2004), who investigated whether the Council is shaped by cleavages based on national or EU-level factors. His findings suggest that the political space of the EU is defined by two scales: a traditional left-right as well as an independence versus integration dimension. Interestingly, he found that presiding countries take the role of an arbitrator. Thus, the "culture of consensus" in the EU Council is based on several pillars which all contribute to the large proportion of unanimity.

This homogeneity in final votes does not allow the evaluation of individual behavior. Therefore, additional data must be considered and two extensions seem most promising. Firstly, one could extend the focus of research to the whole negotiations process. Secondly, the administrative level of negotiations monitored could be extended. The next three paragraphs discuss studies that incorporated such additional information.

Another possibility is to look at interventions during negotiations within the Council of Ministers. Following this approach, Cross (2012) interpreted these interventions as signaling policy positions and as attempts at influencing negotiations. The analysis is based on the records of negotiations of the Council Secretariat. "These documents contain detailed footnotes recording member states' interventions over the course of negotiations. These footnotes identify both the member states making interventions and also the level of negotiation within the Council at which the interventions took place" (Cross,

[139] The monthly summaries of the EU Council meetings can be accessed at `http://www.consilium.europa.eu/documents/legislative-transparency/monthly-summaries-of-council-acts`.

2012, p. 56). The results showed differences between member state intervention behavior which are influenced by the structural characteristics of the policy space.

Hagemann and De Clerck-Sachsse (2007) analyzed the impact of the 2004 enlargement on the internal working processes in the Council. Their data comprised formal statements of Council members following the adoption of a proposal. “These formal statements often consist of a country’s explicit disagreement or reservation with regard to a policy” (Hagemann and De Clerck-Sachsse, 2007, p. 3) and are included in the Council minutes. The authors claimed that they serve a signaling purpose to clarify disagreement with either a single issue or the enacted proposal in general and interpreted them as “a way to avoid policy gridlock through contested voting” (ibid.).

Negotiations are not restricted to the ministerial level alone.[140] Haege (2007) investigated the influence of national bureaucrats on the decision-making in the EU Council of Ministers. Yet, his findings showed that these national officials tend to only decide the less salient and more complex proposals. Hayes-Renshaw et al. (2006) compared voting at the ministerial and official level of the Council. In line with the previously discussed studies the authors found that ministers generally endorse consensus, and that contestation occurs rather at the level of officials. Nearly half the dissent was expressed by single members and most often the cases in question were concerned with “apparently quite technical agricultural issues” (Hayes-Renshaw et al., 2006, p. 5).

To summarize, this section discussed three possible sources of information for an individualized analysis. The suggestion to include other negotiating levels is not possible, since the encoding of NSP would be highly questionable in this case. The expert interviews referred only to ideal positions and salience assessments of member states observed at the ministerial level. A feasible option would be to include either interventions expressed during or formal statements made after the negotiations. Yet, of the proposals covered in the DEU data and affected by NSP only 6 express the required dissent at the individual level.[141] These are too few observations to receive reliable evidence. Maybe the evaluation of NSP on the individual level can be accomplished in future research by including data from the DEU II project (Thomson et al., 2012). However, this data will not contain individual-specific information of NSP. Thus, so far the analysis of individual voting is desirable but not yet available.

3.7 Chapter summary

This chapter offered an assessment of the merits gained when considering NSP in an empirical analysis. It also provided a first glimpse of the necessary expenses. For this pur-

[140] Hayes-Renshaw and Wallace (2006) provided a comprehensive overview on the individual functions and processes of the EU Council of ministers and its development over time.

[141] The corresponding proposals are COD96112, CNS98092, CNS98347, CNS99138, CNS99163 and CNS99235 (cf. Sec. A.5).

pose I used the field of legislative decision-making in the context of EU politics. Models used to assess decision-making rely on utility functions which have to be specified accordingly to the policies in question. When such proposals consist of multiple dimensions the (non)separability of single issues has to be taken into account. Otherwise the utility functions are misspecified and the model produces inaccurate results. This is directly in line with ARGUMENT 1 (established in Sec. 2.5) and confirms HYPOTHESIS 1.

The analysis and coding of the prominent DEU data set provided two main findings: Overall, 55.7% of the issues and 50% of the multidimensional proposals (which constitute one-third of all proposals) in the data set are affected by NSP. Thus, it becomes clear that NSP are in fact a widespread phenomenon in EU politics. Moreover, in the majority of EU law proposals affected by nonseparability the effect is non-reciprocal. The results suggest that the potential bias is most severe in those cases. Thus, the in Sec. 2.4 discussed extension of the NSP concept with respect to reciprocity is strongly encouraged.

Next, using the NSP coding scheme I investigated if a comparison of various models' predictive accuracy depends on the existence of NSP. In my analysis I found clear evidence that the empirical evaluation of competing models of EU legislative politics is biased. The bias arises from the different degrees of restrictions incorporated into the models. These restrictions must rely on a correct specification of actors' utility functions, otherwise they are modeled incorrectly. This is particularly unsatisfactory for those researchers interested in the effect of procedural aspects such as the allocation of agenda power and voting rights. I conclude that this demonstrates the importance of a correct model specification and confirms HYPOTHESIS 2.

This chapter answers my first key question; yes, neglecting the nonseparability of preferences poses a threat to the inference of the corresponding research. Applying simulation techniques I demonstrate that overlooking NSP may have caused a substantial bias in the empirical evaluation of competing models of EU legislative politics. While the effect is not rectified for all models, increasing the level of nonseparability boosts the predictive power of the agenda-setting model. Thus the potential gains of incorporating NSP become clear. Using simulations may also provide a way out to the dilemma of operationalizing nonseparability, although this approach has its own limitations. For example, it is only possible to theorize about and not prove the existence of NSP.

In this chapter I had to talk a lot about limitations of the applied research method. This illustrates all too well that still many aspects are missing in the study of NSP, maybe most evidently in the last paragraph of the analysis. Sec. 3.6 discusses the argument that models based on individual decisions should also model the error term of utility functions at that level. While the idea seems promising, an empirical implementation was not feasible. The DEU data set contains not enough corresponding individual data. One main reason is the low level of contestation in EU Council decisions. Although some studies used additional information to operationalize interventions during Council negotiations the reliability of this data has to be verified by further studies.

The limitations clearly demonstrate why I decided to continue my research on NSP with a laboratory experiment. The high degree of environmental control enabled me to set up the decision problem in the most applicable way, as laboratory experiments secure a high level of internal and construct validity. This chapter discussed multiple models of decision-making and devoted a lot of space for discussions on preference measurement and the importance of various institutional rules. But despite the immense efforts put into the DEU research project I was only able to "assume" NSP based on the expert interviews. These assumptions could only be made in a general fashion as an actor-specific assignment was not feasible. However, a laboratory experiment facilitates to induce NSP for every actor and to observe every single individual decision.

4 The Experiment

I demonstrate very clearly the difficulties of operationalizing NSP in Chap. 3. In particular, I show the limitations of observational data in this context. Thus, using a laboratory experiment is the logical choice for further investigation due to the excellent monitoring capabilities and implementation options. The experimental setting can be configured in detail according to the research interest (cf. Sec. 1.3.2). Thus, NSP are now ensured (because induced) and the investigation can focus on the resulting effects.

The second part of my study consists of four chapters. In those I deal with my second key question, whether and, if so, how nonseparability affects individual and collective behavior. By contrasting several institutional arrangements directly I identify the consequences when a decision is affected by NSP. This follows ARGUMENT 2 and ARGUMENT 3, which are concerned with either sequential or simultaneous decision-making. But I also look for further behavioral patterns to supplement the nonseparability concept.

This chapter starts in Sec. 4.1 with a short overview of important contributions in experimental laboratory research. I pay special attention to the topic of collective decision-making. Next, I explain my experimental design. Sec. 4.2 describes in detail every part of the experiment. The implications of the design choices for the empirical analysis are summarized in Sec. 4.3. Subsequently, I present some descriptive information about the experimental process in Sec. 4.4.

4.1 Experimental research in political science

The basis for laboratory experiments in social science was set in economics (Thye, 2007).[142] The beginning of experimental work in political science is due to the outstanding work of Plott (for an overview on Plotts' work cf. Ortmann, 2003). Since then many excellent contributions have opened the field for further work. Today researchers argue that "lab experiments are a major source of knowledge in the social sciences" (Falk, 2009, p. 535). When conducting experiments, political scientists were mainly concerned with the welfare implications of decision rules like majority and unanimity voting (e.g., Miller and

[142]Palfrey (2009) referred in particular to the work of Vernon Smith (cf. Smith, 1991, for an overview on his work). A testament to his exceptional position as founding father of laboratory research is the statement of the Nobel Prize Committee describing Smith's contribution as "having established laboratory experiments as a tool in empirical economic analysis" (2002 Nobel Prize Announcement, `http://almaz.com/nobel/2002-prizes.html`).

Vanberg, 2014; Sauermann and Glasmann, 2011), majority voting at elections (e.g., Feddersen et al., 2009; McKelvey and Ordeshook, 1982; Morton and Williams, 2011), sophisticated behavior in plurality voting under majority rule (e.g., Felsenthal et al., 1988; Niemi and Frank, 1985), bargaining under majority rule (e.g., Bianco et al., 2008; Diermeier and Morton, 2005; Salant and Goodstein, 1990)), and committee decisions under majority rule (e.g., Coleman and Ostrom, 2009; Fiorina and Plott, 1978).[143]

The focus on majority voting has two reasons. Firstly, majority voting is the most widely spread decision rule for democratic government (Erlenmaier and Gersbach, 2001, p. 2).[144] Secondly, following Rawls (1971), those procedures are considered fair that are acceptable to everybody under the veil of ignorance. Tossing a coin is often accepted as fair procedure because none of the conflicting parties is able to anticipate whether it will be on the winning or on the losing side. Moreover, in infinitely recurring situations coin-flipping produces an equal split, widely accepted to be the only fair division in constant-sum games. Yet in the realm of politics, which is characterized by heterogeneous preferences and binding, non-recurring decisions, neither coin-flipping nor random draws from a lottery are frequently accepted as producing just outcomes.

Instead, social choice theory proves that simple majority voting comes closest to satisfying all criteria of a "fair procedure" (Fey, 2004; May, 1952). This holds even if it is notwithstanding the conflict between procedural justice (Rawls, 1971) and welfare enhancement inherent in majority voting (Arrow, 1950; Risse, 2004). The quality of majority voting with respect to redistributive justice depends, among other things, on the distribution of preferences. Here the advantages of induced, and thus known, preferences enable the experimenter to conduct a detailed and clear cut investigation of the decision-making process.[145]

4.1.1 The rise of the experimental method

There are a multitude of summaries on experimental research in political science, like Kagel and Roth (1995), Croson (1999) or **?**. In addition, newer introductions to political science discuss the research contribution of experiments (e.g., Bernauer et al., 2009, p. 91ff). I do not intend to repeat these volumes but instead focus on a small composition of excellent contributions. Those strongly influenced my decision to use a lab experiment and how to design it. The next sub-section deals in detail with the beginnings of experiments in research on collective decision-making.

[143] For an extensive review please consult the "Cambridge Handbook of Experimental Political Science" (Druckman, 2011).

[144] Any democracy is commonly associated with political equality and majority rule (Saunders, 2010a, p. 2). Yet, most legislative processes in democracies do not use simple majority voting but rather a system of checks and balances, division of power and representative elements (Beitz, 1990, Chap. 4, p. 60).

[145] A central aspect of induced value theory (Smith, 1976) is that by defining the financial incentives the same motivation can be assigned to all subjects. The common payoff structure controls for individual differences (Morton and Williams, 2012, p. 17).

The experimental method has come a long way in political science.[146] For most of the time, experiments were considered inappropriate because this discipline was "limited by the impossibility of experiment. Politics is an observational, not an experimental science" (Lowell, 1910, p. 7). This does not mean that the benefits of the method have not been appreciated. On the contrary, many contributions agreed that "the experimental method is the most nearly ideal method for scientific explanation, but unfortunately it can only rarely be used in political science because of practical and ethical impediments" (Lijphart, 1971, p. 683).

I will not deny that there were also some other perceptions in previous research. For example, Campbell (1969, p. 409) argued that (experimental) administrators "should be ready for an experimental approach to social reform, an approach in which we try out new programs designed to cure specific social problems, in which we learn whether or not these programs are effective, and in which we retain, imitate, modify, or discard them on the basis of apparent effectiveness on the multiple imperfect criteria available." Burgess and Robinson (1969) pointed out that artificial confirmation of a hypothesis in the laboratory (though always possible), was less likely than artificial dis-confirmation; whereas the opposite holds for research in the "natural setting". But these benevolent statements captured by no means the majority opinion.

Even if experiments in social science may never achieve the same degree of "straightforward elegance" (Johnson, 2008, p. XI) as in natural science the critical perception changed step by step.[147] Factors which facilitated this development are the advances in computer technology (e.g., lower costs and better networks)[148], growing research on causal relationships (e.g., Braumöller, 2003) and its difficulties with observational data (e.g., Jeffrey et al., 2006), increasing interest into research questions on underlying assumptions about the nature of political or institutional decision-making, etc.

"From nature to the lab" by Morton and Williams (2010) offered an extensive overview on experimental research and its present state. It contains a multitude of detailed examples from a wide array of scientific research. With a focus on social science in general the authors presented a distinctive approach on how to conduct experimental research. The monograph covers laboratory, survey, and field experimentation. The design of my experiment follows their approach by emphasizing causal inference.[149] Many details of my design apply to recommendations of this book, e.g., the subject recruiting as well as the payment mechanism. The work of Morton and Williams (2010) constitutes a compre-

[146] For an assessment of the history of the experimental method in economics cf. Plott (1991). The analogy of the development in both disciplines is clearly observable, even if the current degree of acceptance differs (McDermott, 2002).

[147] In his magnificent collection of outstanding experiments in natural sciences Johnson (2008, p. XI) referred to the logical simplicity of apparatus and analysis of these experiments as "beauty in the classical sense".

[148] Cf. Humphreys (2004) for a comprehensive discussion on computational methods and their influence on scientific research.

[149] Causal inference was thoroughly discussed by Shadish et al. (2002). The authors focused on theoretical aspects but included also practical advice for implementing an (quasi-)experimental design.

hensive and detailed volume.[150] I follow up with two smaller reviews on the history of experiments in political science. Both of them are shorter and by no means as encompassing. However, both discuss a specific aspect of experimental research I would like to highlight with respect to my own contribution.

Faas and Huber (2010) entitled the increasing application of laboratory experiments with the expression "From Wallflower to Mainstream". The article reviewed the state of experimental research done in the fields of elections, public opinion, public goods, collective action, social trust, legislative bargaining and decision-making. It also contained a more detailed section on field experiments which studies voter mobilization. The authors argued that experimental research is on the rise, as an increasing number of related journal articles, books, and conferences show us. Yet, they also noted that in particular in German political science the use of experimental methods is still rather uncommon.[151] As a reason for this the authors emphasize the missing scientific training with respect to experimental methods in political science curriculum at German universities.[152] Despite its small-scale proliferation in Germany I choose the experimental approach because I agree with the advantages of the experimental method discussed by the authors which clearly demonstrate the added value of this method.[153]

Political science has often been influenced by associated fields of studies with regard to research methods (Druckman and Lupia, 2006). McDermott (2002) compared the historical and cultural differences when conducting experiments in two neighboring scientific fields, economics and social psychology (for an overview cf. Martin, 2008). Although both disciplines use experiments widely, their approaches differ with respect to many fundamental basics.[154] Most prominent are the attitude towards deception or the monetary payment of subjects. Bloomfield et al. (2009, p. 14ff) argued that "many experiments in economics are not actually experiments - they are demonstrations." A typical psychology experiment manipulates a single variable while holding all other aspects constant. Accordingly, Shadish et al. (2002, p. 7) characterized experiments as to "explore the effects of things that can be manipulated". But in economics testing theories is also common, which can be done without any manipulation. As my research framework resembles rational choice I followed the tradition of experimental economics.[155] Thus, subjects

[150] With Morton and Williams (2012) the authors also provided a shorter and apt summary of the usage of experimental methods in political economics.

[151] Of course, there are also attempts in Germany to promote the experimental method further. For example, the working group for action and decision theory of the German association for political science devoted its 2012 yearbook to it (cf. Bräuninger et al., 2012).

[152] In addition, in Germany political science students usually do not complement their education with additional psychological, sociological or economic methods courses, in which they may learn experimental procedures (Faas and Huber, 2010, p. 724).

[153] I explain my decision to investigate NSP by conducting a laboratory experiment in detail in Sec. 1.3.2.

[154] Cf. Croson (2005, p. 131) for a in depth comparison of experiments in economics and psychology with respect to "incentives, context, subject pools, deception, experimental details and data analysis".

[155] Cf. Guala (2005) for a comprehensive discussion on the principles of experimental inference and the scope and limitations of the laboratory in experimental economics. In addition, cf. Bergstrom and Miller (2000) who focused on topics of microeconomics.

received a monetary payment (depending on their performance in the experiment) and no crucial or necessary information was withheld.

All reviews have in common that they take a great deal of their space and devote it to explain the details of the experimental method. They define the basic principles of causality, randomization and validity as well as discuss pros and cons of different types of experiments conducted in either the laboratory, field or as a survey. Of course, all authors have done experimental research themselves and thus can be expected to advertise this method. But in every volume potential errors, traps, disadvantages and limitations of the method are also discussed. Overall, I agree with the pointed appeal of Kinder and Palfrey (1991): "An Experimental Political Science? Yes, an Experimental Political Science!" Yet, not without adding that, while experimental methods have much to offer political science research, no method can ever be suited to fit all research questions, however flexible and universal the approach might be.[156]

4.1.2 The first experiments on collective decision-making

Collective voting behavior is and was an area of public choice characterized by a considerable amount of disagreement (cf. Clinton and Meirowitz, 2004). This gave rise to an extremely "fruitful area for experimental research" (Anderson and Holt, 2002, p. 4). The beginnings go back to simple paper and pencil experiments which focused on the effect of different decision rules. Here, Fiorina and Plott (1978) used a plain design and put five-person committees in front of a blackboard and introduced a single point in the two-dimensional space as the SQ. "Any subject could propose an amendment [...] to the motion on the floor. If it passed [...] the amendment became the new motion on the floor and the process continued" (Fiorina and Plott, 1978, p. 577). At any point of the experiment a majority could end the debate. Thus, each "committee pushed a point around the blackboard until a majority voted to quit and go home" (Fiorina and Plott, 1978, p. 577).

The canonical experiment of Fiorina and Plott (1978) has been advanced with respect to alternative procedural rules and sequential decision-making by Kormendi and Plott (1982) as well as Eavey and Miller (1984, 1995). All together, these contributions compared for committee experiments "the predictions of a variety of models drawn from Economics, Sociology, Political Science and Game Theory [...] to the experimental results" (Fiorina and Plott, 1978, p. 575). The later studies paid more attention to general aspects of agenda control and information levels. Mostly, the committees were made up of an odd number of participants and voted under simple majority rule on monetary payoff schemes which resembled closely a "surface plot" as shown in Fig. 4.1.

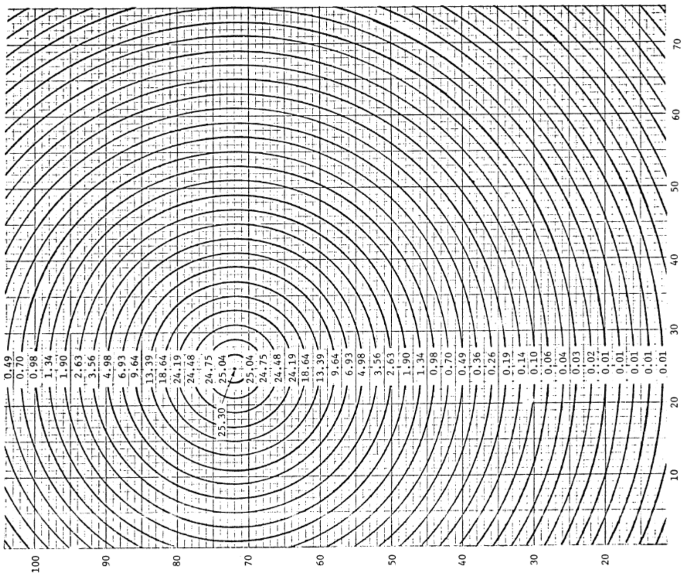
These experiments found that equilibrium concepts such as the core predict the observed behavior fairly well (Tullock, 1981), but are blurred by a level of uncertainty which varies

[156] Moses and Knutsen (2007) offered a comprehensive overview of methodology in social sciences. The volume emphasized the debate between positivist and constructivist approaches and stands representative for the methodological pluralism of the discipline.

Figure 4.1: Payoff scheme

EXPLANATORY NOTE

The figure shows a payoff scheme handed out to every subject in the study of Fiorina and Plott (1978, p. 596). Their experimental design was straightforward. Each committee was asked to select a single point in a two-dimensional decision space (x_1,x_2) on a blackboard by majority rule. Each subject was assigned a payoff function defined over ordered pairs of the coordinate system. Those pairs represented the amount of money the subjects would receive if that point was chosen as a collective outcome. From the scheme every subject could infer its payoff for each point of the decision space.



across designs. More advanced concepts for dealing with random errors in uncertain environments included the selection set (Salant and Goodstein, 1990) and, more recently, the quantal response equilibrium (QRE; McKelvey and Palfrey, 1995, 1998). If, however, the core is empty as it would be the case in a multidimensional, simple majority setting (Berl et al., 1976), the outcomes are located close to the center of the policy space, best approximated by concepts such as the yolk (Plott, 1967) or the uncovered set (Bianco et al., 2004; Fishburn, 1977).[157] In my analysis I used a two-way random error component model (cf. Sec. 5.1) as well as a random utility mixture model (cf. Sec. 6.3) to account for uncertainty and errors in the estimations. This also included a probabilistic extension of the core which enabled a far more accurate analysis then a deterministic approach.

4.2 Experimental design

My experiment falls into the category of investigating "Formal Theory" (Morton and Williams, 2010, p. 145ff). It resembles a "stress test" (Schram, 2005, p. 234) for theories of collective as well as individual decision-making (for an overview on such theories cf. Davis and Holt, 1993). Morton and Williams (2010, p. 151) defined such a stress test as "when the researcher chooses to allow for one or more of theoretical assumptions to be violated or purposefully investigates situations in which the researcher is uncertain as to whether theoretical assumptions hold".

[157] For a recent overview cf. Bianco et al. (2008).

I adapted the general experimental environment presented by Sauermann and Kaiser (henceforth S&K, 2010). These authors conducted a series of laboratory experiments designed to elicit the relevance of social preferences for majority decisions. They concluded that fairness motivates majority decisions. In my experiment I studied decision-making in committees under majority rule. Subjects voted simultaneously for one out of several alternatives. If neither of the alternatives got a majority, a new ballot was held. This process continued until one of the alternatives received a majority of votes which marked the completion of the voting process. Subjects were provided with cardinal information, i.e., they knew their own and other players' payoffs and they were informed about other players' previous votes.

During the experiment I altered the procedural rules in a way that subjects were confronted with separable as well as nonseparable decision situations. Between these modifications I kept other factors as subjects' preferences and information condition constant. I was interested in the way subjects adjust to the nonseparability of the decision-making situation and if the observed outcome varies with alternations of the procedural rule. This included behavior at the collective as well as at the individual level. I favored a more realistic framework and used multiple variations of preference distributions.

It is first and foremost important that subjects understand the game they are playing. Chou et al. (2009) emphasized the importance of game-form recognition very distinctly. Only under this precondition experimenters can set up an experimental design that is highly applicable to the research interest. To secure the game-form recognition I also conducted a pilot session as test run and a post-experiment survey (cf. Chap. 7).

4.2.1 Procedure

All sessions were conducted in the experimental laboratory managed by the Department of Economics at the University of Heidelberg[158]. Subjects were recruited using the ORSEE software (Online recruitment system for economic experiments, Greiner, 2004) out of the subject pool of the laboratory. I implemented my design using the z-Tree (Zurich Toolbox for Readymade Economic Experiments) software developed by Urs Fischbacher (2007).

The participants entered the laboratory one by one and were assigned to cubicals by drawing lots. The instructions, which are included in Sec. A.7[159], were handed out to the subjects and read aloud by a supervisor.[160] Questions were answered privately.

After that the subjects went through the rules on their screen again and had to answer some questions correctly in order to prove their understanding of the upcoming game. Special emphasis was placed on ensuring that participants understood the key aspects of

[158]Homepage of the laboratory: `http://www.uni-heidelberg.de/fakultaeten/wiso/awi/forschung/awiexplab.html`.

[159]Please note that the instructions are only available in German.

[160]The instructor was the same person throughout all the experimental sessions.

the experiment, such as their decision-making competences, the timing of the ballots, the set of feasible alternatives, etc. Also, the basic features of the PC setting as, e.g., the screen arrangement, the payoff table scheme and the entering field for their vote were part of the test. The experiment did not start before all subjects had passed this mandatory test.

Next, the participants took part in multiple rounds of my voting experiment (cf. Sec. 4.2.5). After the conclusion of all ballots the participants answered a post-experiment survey at their PC workstation (cf. Chap. 7). Upon completion, they were asked to leave the room and to stay in the waiting area. After everyone had gone out, the participants were individually asked back into the experimental laboratory and received their payoffs.

4.2.2 Anonymity and non-communication

During the whole experiment, communication between the subjects was forbidden. Previous studies found that communication enforced the relevance of social norms. This was evident, e.g., in form of "universalism" in Miller and Oppenheimer (1982) or "collective resistance against transgression" in Cason and Mui (2007). Yet, also under an anonymous setting the participants not only maximize their individual payoffs (van de Kragt et al., 1983). Moral motivations may be enhanced by communication, but they do not dependent on it (Fehr et al., 2002a).[161]

In addition to (active) communication "silent identification" (e.g., the disclosure of name and place of birth) has also been found to increase solidarity between subjects (e.g., Bohnet and Frey, 1999). More specifically, in their public good experiment Andreoni and Petrie (2004) experienced less free riding and, if identification was combined with information on others behavior, increasing contributions to a common public good.[162] I followed none of these approaches. My design excluded pre-play communication and guarantees all participants absolute anonymity. Other players were indicated only as *Player 1*, *Player 2*, *Player* 3, etc.

Anonymity also applied for the payment process. Payoffs were handed out when only the receiving subject was in the room. This leaves only the experimenter as person the subjects could have interacted with. This possibility for distortion is called experimenter demand effect (EDE).[163] Such an effect occurs when subjects change their behavior in

[161] In their public good experiments van de Kragt et al. (1983) observed significant fewer realizations of the public good when communication was not allowed. Nevertheless, subject's contributions were still sufficient in 65% of the trials.

[162] The identification and monitoring of others is strongly associated with experimental research on group identity. In both cases the decision of the subjects may (partly) be influenced by a feeling of solidarity on the basis of a common ground. The difference is that group membership mostly rests on an "artificial" aspects (e.g., which of two paintings they like best, Chen and Li, 2009, p. 436) and identification on "true" personal characteristics (e.g., gender). Chen and Li (2009) provided a comprehensive overview for literature on this topic.

[163] Due to the architecture of the laboratory a "double-blind" implementation was not feasible. Such an experimental procedure would allow neither the subjects of the experiment nor the persons conducting the experiment to know the critical aspects of the experiment (Shuttleworth, 2008). Cf. Cox et al. (2008, p. 18-20) for more information on the comparative advantages of single-blind and double-blind protocols.

an experiment due to cues about what constitutes appropriate behavior and would be expected (or "demanded") from them (Zizzo, 2010, p. 2). Barmettler et al. (2011) investigated this potential contamination of experimental data as many experiments do not provide anonymity between experimenter and subjects. The authors varied the degree of anonymity in common laboratory setups but found no significant EDE on subjects' behavior. This is in line with Frank (1998) who tested if subjects care about an experimenter's welfare. Using ultimatum bargaining experiments he found no such effect. Zizzo (2010) showed that EDE can be a problem when they are correlated with the true experimental objectives; but he also admitted that "given the trade-offs implicit in designing and running an experiment, researchers may decide to accept the risk of an EDE" (Zizzo, 2010, p. 28). So far the literature provides enough evidence that the experimenter-subject contact should constitute no distorting factor. However, a post-experimental survey gave me the opportunity to look for indications of EDE.

4.2.3 Payoff table

The number of points each player earns in case a given alternative was selected was specified in neutrally labeled payoff tables as, shown in Fig. 4.2. Here, every participant (player 1 to 6) is assigned a certain number of points in each case. A total of nine alternatives were available. These alternatives formed a combination of columns (letters A, B and C) and rows (numbers 1, 2 and 3). Thus, $A1$, $A2$, $A3$, $B1$, $B2$, $B3$, $C1$, $C2$ and $C3$ constituted the nine possible alternatives.

Figure 4.2: Example of a payoff table

EXPLANATORY NOTE

The figure shows a payoff table as presented to the participants in my laboratory experiment. The two-dimensional design enables the implementation of NSP when separating the issues. Then, a decision on each single issue influences the preferences for the outcome of the other (Strom, 1990, p. 57). As long as no restrictions are imposed or (partial) decisions are made, every individual prefers one of the given alternatives as their first choice (i.e., unconditional first preference). For example, assuming rational and self-interested actors, this is $\{B2\}$ for player 5 with 45 points (i.e., his maximum score). By altering the options or fixing a partial decision an individual's preference for the preferred outcome changes also in the second dimension. For example, in case options $\{A\}$ and $\{B\}$ are excluded, player 5 is preferring row 1 (with a payoff of 26) instead of row 2 (now with a payoff of 12). This makes $\{C1\}$ the conditional first preference.

		A	**B**	**C**
1	Player 1	7	2	1
	Player 2	28	8	41
	Player 3	27	12	24
	Player 4	22	44	4
	Player 5	4	2	26
	Player 6	6	9	20
2	Player 1	58	35	5
	Player 2	2	11	3
	Player 3	10	4	45
	Player 4	19	1	8
	Player 5	8	45	12
	Player 6	11	10	28
3	Player 1	8	21	18
	Player 2	10	32	23
	Player 3	10	18	15
	Player 4	17	13	11
	Player 5	10	6	20
	Player 6	32	17	4

I build my research on a solid foundation of previous experimental work on collective decision-making (cf. Sec. 4.1). Contrary to these contributions the preferences were not generated by monetary payment schemes or geometric representation (cf. Fig. 4.1) but payoff tables. This followed two reasons. Firstly, the change is justified by the results of the prior work which often were based on kind of "vague" statements: committee decisions "continue to cluster around an alternative"[164] (Fiorina and Plott, 1978, p. 579) or "the influence observed to date conforms closely to that predicted by the core" (Kormendi and Plott, 1982, p. 189). S&K (2010, p. 669) summarized these findings with "a general tendency toward the core point". I do not question the overall conclusions these authors draw from their data, as their contributions are too elaborated and numerous. The core seems to be a good approximation for committee behavior. Nevertheless, the usage of payoff tables provided more clear-cut results than the spatial proximity information. An alternative is selected or not, enabling an analysis with higher discriminatory power.

Secondly, the experimental data of Diermeier and Gailmard (2006) suggested that participants evaluate endogenously and exogenously generated inequality differently.[165] Accordingly, Mertins (2008) pointed out the effect of various procedures and the importance of procedural fairness. Endogenous inequality results from decisions made by the participants in the laboratory while exogenous inequality resorts to intentional modifications the experimenter makes before the beginning of the experiment (e.g., giving one subject an initial endowment of 50 while another subject receives only one of 5). Rabin (1993, p. 1296) argued that "people's notions of fairness are heavily influenced by the SQ and other reference points." The experimental studies on redistribution of Rutstroem and Williams (2000) and Sutter (2002) found that subjects' voting behavior depended systematically on their relative position to the SQ.[166] Thus, in line with the design of S&K (2010), but contrary to prior contributions[167], I did not specify a SQ ante (i.e., select one alternative as default option) because it might bias subjects' behavior (cf. Samuelson and Zeckhauser, 1988).

In line with the experiments discussed in Sec. 4.1.2 the dimensions of the payoff table were labeled with "neutral" numbers and letters. Lacy (2001b) showed that conditional preferences exist for many policy areas. Therefore, framing[168] could be used to induce

[164]Continue to cluster stands for the frequent occurrence of outcomes in the near vicinity of an alternative.

[165]Maier-Rigaud and Apesteguia (2003) identified another interesting "endogenously vs. exogenously" effect. The authors varied if a prisoner dilemma is automatically assigned to offering the subjects the possibility of choosing between two different representations of the same dilemma. They found significantly more cooperation when the game could be chosen.

[166]Colomer (2001, p. 10) stated in his evaluation of different voting rules that "the social efficiency of the outcome is highly dependent on the SQ, the distribution of bargaining costs among voters, and the agenda setter's maneuvering." Thus, the implemented SQ might not determine the final outcome alone but it will exert significant influence.

[167]For example, in the two-dimensional setting of Fiorina and Plott (1978, p. 577) "each committee began at the point (200, 150). That is, the status quo in each experiment was the extreme northeast point in the issue space".

[168]Framing occurs when different, but logically equivalent, phrases cause an individual to alter their preferences (Tversky and Kahneman, 1986).

conditionality (Tversky and Kahneman, 1981). however, this would bring about other problems. First, as with salience, the strength and direction of NSP varies for every individual. The empirical analysis in Chap. 3 clearly shows how much the salience of a single issue can vary across actors. In order to analyze the decision on a framed question empirically, one would have to measure these individual values.[169] In the context of the laboratory this could have been achieved by conducting surveys on participants' perception of the experimental decision problems. Nevertheless, the downside of such an approach would be that the identification of individual salience and NSP would, even in the best case, only become clear ex-post.

A second problem with using framing is that subjects themselves introduce a number of latent dimensions into the decision the experimenter wants them to make (Halfpenny and Taylor, 1973). As these dimensions are not homogeneous across subjects, it is not possible to identify all of them even ex-post.[170] To avoid these problems and because my experiment was the first which focuses explicitly on the aspect of nonseparability, I chose to apply a neutral context. This generic environment ensured the highest possible level of control over the motivation of the subjects (cf. Morton and Williams, 2012, p. 22).[171]

4.2.4 Multiple tables and their characteristics

I used different payoff tables in the experiment. The aim was to observe decision-making under a variety of problems. This resembles reality with its multitude of everyday decision situations better than the implementation of just one fixed payoff scheme. Sec. A.8 contains a list of all tables used in the experiment. The tables differ with respect to two important characteristics: i) if an equilibrium prediction exists and ii) if the alternative's overall sum varies or stays constant within a table.[172] These variations are not the treatment of the experiment (cf. Sec. 4.2.6), but rather allowed me to study the treatment effect under different constellations (cf. Sec. 5.4).

[169] For example, the DEU research project (cf. Sec. 3.2) used in-depth expert interviews.

[170] Halfpenny and Taylor (1973) conducted a series of committee experiments. In an attempt to add more interest and realism to the experimental situation they told their subjects that they were facing a decision for the location (inner city vs. suburb) and fittings of a new office block. From comments made by their subjects during the pilot session "it became apparent that they were, in effect, introducing an unspecified number of additional dimensions, e.g., whether to build in the Green Belt, or the availability of commuter trains." (Halfpenny and Taylor, 1973, p. 28) Thus, they choose to conduct their experiment with labeling the dimensions "X" and "Y".

[171] In general, the question for using neutral or loaded instruction in either payoff tables or payment schemes must depend on the goal of the research. Eckel and Grossman (1996) emphasized that the importance of social and psychological factors can only be studied by abandoning abstraction (at least to some extent). On the other hand, to assume the effect of loaded instructions as incontrovertible may be wrong. Quiet some empirical studies found no effects even if "the underlying context is heavily loaded" (Abbink and Hennig-Schmidt, 2006, p. 103). For further information on priming effects cf. Ortmann (2005) and Ortman and Gigerenzer (2000).

[172] The experiment inhibited a certain intrinsic aspect. In addition to theory driven hypotheses I was also interested in emerging patterns of behavior I could not anticipate before. Such observations are much more likely and substantiated when looking on a multitude of decision situations.

Following social choice theory (e.g., Fishburn, 1973, 1977) the appropriate baseline equilibrium for the committee decision-making is the concept of the core (cf. Sec. 2.4.2).[173] This is the same solution concept as in S&K (2010). Previous work suggests the validity of this approach, e.g., the experiments discussed in Sec. 4.1.2. The core alternative is characterized as a unique cooperative solution under the assumption of rationally acting individuals. More generally, the core is defined as the alternative for which no single player or subgroup has an incentive to leave the coalition supporting it (Peleg and Sudhoelter, 2003, Chap. 3 and 12). Hence, the core alternative beats all other alternatives in pair-wise comparison (analogous to Berl et al., 1976, p. 468).

The core is a comprehensible and simple concept. Nevertheless, it is not limited to a specific institutional rule. Miller et al. (1996) conducted experiments to test whether individuals partitioned into two chambers using simple majority voting do, in fact, agree on the "bicameral core"[174]. They found that the chambers almost always chose the outcome the formal theory predicted. This is a very useful insight as it allowed me to operationalize different decision-making mechanisms together with their respective equilibrium concept (cf. Sec. 4.2.6).

With a view to the applied stopping rule this choice may be disputable.[175] Subjects voted simultaneously, and if an alternative received a majority of votes this marked the completion of the voting process. There never was a pair-wise vote between two alternatives. In fact, this comparison had to be made by the subjects while considering their decision. This stopping rule was also used by S&K, (2010), but deviates from the procedures used in previous studies (cf. Sec. 4.1.2). Yet, although those contributions do not agree on one single best practice but cover a whole range of stopping rules as, e.g., motions to adjourn (Fiorina and Plott, 1978; Bottom et al., 2000) or 5-times pairwise winner requirements (Halfpenny and Taylor, 1973).

Even the underlying idea of my rule can be found in this wide variety. Kormendi and Plott (1982) appointed one of their subjects' convener with the sole ability to make proposals the group could ratify. Other participants were limited to make suggestion for proposals. While in their setting "the convener may propose any option he/she wants or he/she can refuse to propose any option if he/she so desires" (Kormendi and Plott, 1982, p. 192), another aspect is in accordance with my implementation. Once ratified (i.e., proposed by the convener and accepted by a majority), an option it is no longer contestable but immediately final.

[173] Kramer (1972, p. 169) stated that "any n-person game, can be analyzed from either a cooperative or non-cooperative point of view". The decision for either depends on whether one judges "that the terms of the original agreement can be enforced" (ibid.). I discuss this aspect in Sec. 5.2.1 in detail.

[174] Miller et al. (1996, p. 87) defined the bicameral core as "the core, which will depend on the way in which a given set of voters is divided into two chambers, a majority of each chamber being required to pass legislation."

[175] In addition to the stopping rule other factor my influence the suitability of the core concept as well. McKelvey and Ordeshook (1981) analyzed spatial majority voting games. Their results suggest that the performance of the core can be affected by several different aspects; e.g., the structure of the alternative space or the dominance relation underlying the social ordering.

The main difference between my design and the previous contributions is the information level of the subjects. My experiment is characterized by complete information of one's own and other subjects' payoffs (cf. Sec. 4.2.3). This should enable the participants to make the necessary considerations. Contrary to my approach, other designs (e.g., Fiorina and Plott, 1978; Eavey and Miller, 1984) limited the actual level of information to, e.g., the ordinal value of alternatives or the ideal positions but not the preference distribution of other subjects (S&K, 2010, p. 669).

Previous experiments without a stopping rule have shown that once the choice had converged on the core alternative a deviation is highly unlikely (S&K, 2010, p. 668ff). I reviewed these findings in a pre-test session of my experimental design. Here, the collective decision was not immediately locked in after enough subjects of a group had agreed on one alternative. Subjects were informed that the threshold was reached and which alternative constituted the collective decision, but they were also asked to vote again.[176] In this pre-test 18 participants were in the laboratory and were separated into three groups of six players. The voting procedure of one round did not end until all three groups had reached a collective agreement. Thus, some groups had to vote no less than five more times. I conducted three rounds in which subjects voted in the group of six and further three rounds in which subjects voted in groups of three players.[177] Overall, I collected 18 collective decisions. In all rounds and under all procedures the first reached collective decision was not altered. Admittedly, I found some variance in voting behavior across groups; i.e., different groups agreed on different alternatives and subjects exposed to the same decision problem voted differently. But not once was the first reached collective decision revised.

In addition, work on voting under plurality rule has shown that strategic voting pushes the collective decision towards a stable equilibrium (Feddersen, 1992). Here, many contributions investigated party-positioning games (for an extensive review on this topic cf. Coughlin, 1990a,b). Other models by Cox (1997) and Palfrey (1989) assumed nonstrategic parties but strategic voters when looking at Duverger's Law (Duverger, 1954). Palfrey (1989) argued that as the size of the electorate increases the rational support for all parties, except for the top two, goes asymptotically toward zero. Fey (1997) investigated the "wasted vote" phenomenon and how voters can coordinate. He showed that in a Bayesian game model of strategic voting non-convergence is possible, but this constitutes only an extreme case.

The second main difference between the payoff tables lies in the alternative's overall sum. The experiment contained constant-sum as well as non-constant-sum tables. In a constant-sum payoff table the overall sum of an alternative (the sum of points of all six players) stays constant. All nine alternatives provide the identical amount of points. In

[176] Subjects were also still provided with information about other players' previous votes.

[177] These modifications refer to voting procedures I use as treatment in my experiment. I discuss the procedures and their differences in detail in Sec. 4.2.6.

non-constant-sum tables the alternative's overall sum varies between alternatives. This implemented two different decision situations for the participants.

In general, experiments can be classified according to their potential for conflict. This is based on the overall experimental setting and its inherent dynamics. For example, compared to a public good game the use of a bargaining setting implements a distinctly higher level of conflict. Bargaining focuses not on achieving a goal together but on distributing a collective resource; a task that clearly comprises a certain amount of disagreement. Hinich and Munger (1997, p. 7) stressed that such "disagreement tests collective choice mechanisms; conflict strains the ties that gather a group of individuals into a society." My design resembled no standard bargaining procedure. In such a setting a first actor proposes a distribution and a second actor either accepts or rejects the proposal (cf. Kagel and Roth, 1995, Chap. 4). Yet, I followed a similar intention by using constant-sum as well as non-constant-sum payoff tables.

A constant-sum table leaves the subjects with one single question: how should they split the exogenously defined amount of points. Comparing two alternatives, every gain of one player is mirrored by a loss of at least one other player. The focus rests on distribution and constrains the tactical opportunities of the subjects. Because of this limitation "zero sum games can be regarded as the branch of game theory with the most solid theoretical foundations" (Palacios-Huerta and Volij, 2008, p. 3). Non-constant-sum tables bring another aspect into play, the overall sum of points. This may be seen as a measure for social welfare or an alternative's effectiveness. This sum differs between the given alternatives. Only the possible range for the overall sum is exogenously given. The committee decides about the exact amount endogenously. Thus, subjects have to deal with a trade-off between distribution and allocation.[178]

This characteristic is best met by the concept of game harmony put forward by Zizzo and Tan (2009).[179] They described a subject's perception of the experimental game as decisive for their decision "how to play". Game harmony is then "a generic game property that describes how harmonious (non-conflictual) or disharmonious (conflictual) the interests of players are, as embodied in the payoffs" (Zizzo and Tan, 2009, p. 3). Harmony measures can be calculated solely based on the payoff matrix. My two table set-ups were located at the extremes of the game harmony scale.[180] Constant sum-games are games

[178]Not all findings correspond with this view. Bottom and Paese (1997, p. 1920) pointed out that even in bargaining situations with the possibility for joint gains the "negotiators often behave as though their own interests and their counterpart's were completely opposed. Rather than engaging in information sharing, exploration of options, and joint problem solving, negotiators exchange threats, make ultimatums, and rely on other tactics that might be more appropriate in a zero-sum game." Neale and Bazerman (1991, p. 61) denoted this as the "myth of the fixed pie" where participants do not perceive a possible common progress.

[179]Contrary to my experiment the study by Zizzo and Tan (2009) focused on the relationship between game harmony and cooperation in static games. In their experiment players move simultaneously and the game is played only once.

[180]Zizzo and Tan (2009, p. 4) pointed out that "in the coordination game there is perfect harmony of interests between the players: the only problem is one, indeed, of coordination. In the constant-sum game, the gain of a player is the loss of another, which means that there is perfect disharmony of interests."

of pure disharmony, as every subject's gain is another subject's loss. On the other hand, games with varying payoff-sums still require coordination. But cooperation even is in the interest of the most self-interested subject.[181]

The payoff tables were assigned randomly in every experimental session. Therefore I had to account for (statistical) effects which resulted out of the different usage of tables. Without this calibration the results would be biased. For example, consider the overall sum of the alternatives. They range in payoff table 17 from 104 points to 145 points. Looking at payoff table 19, the range extends from 83 points to 91 points. When interested in the average number of points collected, I have to take the possible range into account. If one were to use frequently table 17 and less often table 19 one would end up, everything else equal, with a higher absolute average of points. An effect based solely on the payoff table differences. Thus, I abstained from using absolute values when interpreting the results in the empirical analysis but employed relative judgments. The specific normalization is discussed in the corresponding section of Chap. 5.

4.2.5 Multiple rounds

Every experimental session lasted multiple rounds. In every round I varied two aspects. Firstly, committee members were matched randomly out of 18 participants.[182] This prevented subjects from building a reputation between rounds.[183] Secondly, the payoff tables were assigned to each round by a random draw. This draw took place in every experimental session without replacement. Thus, every table could be played only once during a session. Yet, no session lasted long enough to use all the tables.[184] Both "random" facts were stated to the subjects at the beginning of the experiment.

Subjects were instructed that the experiment would continue for several rounds and that there was no predetermined upper limit to the number of rounds. Nevertheless, subjects may have guessed the maximum number of rounds from the overall time schedule. The recruiting email sent to all members of the subject pool, informed participants about how long they would be required to stay. Still, subjects were not informed or aware of whether a certain round was the final ballot of the experiment.

Since my experiment extends over several rounds I have to consider possible effects over time. Following Crawford and Haller (1990) as well as Roth and Erev (1995) it is reasonable to assume that subjects changed their voting patterns after they got acquainted to

[181] Dixit et al. (2009, Sec. 2.B, p. 21ff) classified games in a similar fashion. They focus on if "player's interests [are] in total conflict, or is there some commonality?"

[182] The matching followed the stranger matching procedure which resembles "a random matching, i.e., in every period, the group is determined by the computer's random generator" (z-Tree tutorial, source: `www.iew.uzh.ch/ztree/ztree21tutorial.pdf`, version June 3, 2002).

[183] As I excluded reputation building through my design I will not discuss its effects in detail. For comprehensive assessment please cf. Bolton and Ockenfels (2005) and Levin (2009, Chap. 4 and 13).

[184] The longest session lasted for 2:45h and extended over 16 rounds.

the voting game.[185] If this happens, the analysis would miss this behavior if only focused on average values over all rounds. Yet, such a change of behavior is no axiomatic statue. Herzberg and Wilson (1988) found no support of a change in behavior over rounds when they conducted experiments on the proportion of sophisticated and sincere voting. Thus, experience alone may not be enough to change voting behavior. Nevertheless, in my empirical analysis I accounted for this by depicting not only average values but by testing all data for possible trends over the rounds.[186]

4.2.6 Treatment

The experiment focused on problems related to economies of scope, namely the separation of decision-making competences over policy proposals characterized by NSP.[187] This was secured by the large extent of environment control in the laboratory which also enables a high level of internal and construct validity (Morton and Williams, 2010, p. 192).

My payoff tables have a two-dimensional structure (cf. Fig. 4.2). However, this does not automatically results in nonseparable decision problem. A selection of nine alternatives from which one can be unambiguously selected may just as well be shown in either a 3×3 or 1×9 table. The form of presentation (alone) is not decisive[188] because the payoff tables are not assembled in a way that adjacent fields are related. In the spatial presentation of payment schemes (cf. Fig. 4.1) the proximity of any point to the ideal position is crucial; and for nearby points (per definition) similar values apply. Yet, in my payoff table $C2$ needs not to be more alike to $C3$ then $A1$ is to $C3$. The spatial principle of continuity is not found in the tables. The necessary key element for nonseparability is the split into multiple dimensions and the separation of the associated decisions.

Therefore, my treatment is a change of voting rules which the subjects had to follow.[189] The intervention varied the decision-making with respect to two aspects. Firstly, the voting competence of the participants was altered. The subjects were either put all together for a common task or separated into groups with their own responsibilities when casting their votes. And secondly, the voting sequence of the procedure was modified. The ballots were either held all at the same time or successively, one after the other.

[185] A round in my experiment did not only consist of one ballot or a single cast vote. Subjects voted for their preferred alternative in every ballot in every round. Thus, in every round they were involved in a dynamic coordination game with the other subjects of their group. In the next round the composition of their group was changed. As the matching was set randomly they could not reason about the identity of other players now being part of their group. Yet, of course each subject had experienced a variety of reactions, successes or disappointments in previous rounds. Thus, in every new round subjects ought to adjust their voting choices, according to their experiences made during the actual experiment.

[186] For this purpose I used nonparametric statistics (cf. Sec. 5.2) instead of counterbalancing, as this would require the assumption of a linear effect (Martin, 2008, p. 156ff).

[187] By assuming that decision makers are i) policy seekers and ii) randomly assigned to each policy area I was able to disregard the potential for bureaucratic drift (cf. Sec. 2.1) and set a more narrow focus.

[188] The labels of the nine alternatives, regardless of whether $A1$-$C3$, $A1$-$A9$ or A-I, do not matter.

[189] This is not a novel but validated approach when assessing the influence of organizational aspects. Already Shepsle (1979) compared various institutional settings as the division of labor in committee systems, specialization arrangements or the monitoring of sub-level units.

Altogether, the experiment used simultaneous, sequential, pooled and delegated decision-making. Tab. 4.1 gives an overview of the specifications. The experiment comprised three different procedures I refer to as *pooling*, *simultaneous delegation* and *sequential delegation*. All procedures followed the overall design in that if no alternative was chosen by a majority, members received information on how the other players voted[190] and the game continues with a new voting. Ballots were repeated until an alternative receives the necessary votes. After the group had made its decision, all players were assigned their corresponding number of points regardless of whether they had voted for or against this alternative themselves.

Table 4.1: Treatment characteristics

		VOTING COMPETENCE	
		pooled competence	delegated competence
VOTING SEQUENCE	simultaneous voting	pooling	simultaneous delegation
	sequential voting	-	sequential delegation

POOLING

Under pooling (henceforth POL) the group of six subjects voted for either of nine alternatives ($A1$-$C3$). All players cast their vote simultaneously. If neither of the alternatives got at least the qualified majority of four out of six votes, the information on the voting behavior of the group was shown and following a new ballot was held. This process continued until one of the alternatives received the required majority which marked the completion of one round.

This voting procedure is not a decision situation affected by nonseparability. All players decide simultaneously about a specific and unique alternative to be chosen. There are no restraints for the subjects when they cast their votes, nor any temporal sequence which forces a separation into individual issues and conditional adaptations.

SIMULTANEOUS DELEGATION

Under simultaneous delegation (henceforth SIM) two delegations consisting of three players each decided simultaneously over one table with nine alternatives. Each delegate could only vote for a subset of the alternatives, either column or row. Subjects were randomly assigned to each delegation. To reach a decision at least two of the three subjects of a delegation had to vote for the same subset. Within each delegation votes were cast simultaneously. If no single subset got at least the qualified majority, the information on the voting behavior of the other delegation members was shown and the voting continued.

Delegates were not informed about the voting behavior in the other delegation until both delegations had reached a decision. The intersecting of both subsets determined the chosen alternative. For example, one delegation decides on the column (A, B or C) and

[190] In the lower left corner of the decision screen the other players and one's own previous voting behavior (in the current round) is displayed.

chooses set $\{A\}$. The other delegation votes on the row (1, 2 or 3) and chooses set $\{3\}$; then the resulting outcome for all six players is alternative $\{A3\}$.

SEQUENTIAL DELEGATION

Under sequential delegation (henceforth SEQ) two delegations consisting of three players each decided sequentially over one table with nine alternatives. Each delegate could only vote for a subset of the alternatives, either column or row. Subjects were randomly assigned to each delegation. To reach a decision at least two of the three subjects of a delegation had to vote for the same subset. Within each delegation votes were cast simultaneously. If n o single subset got at least the qualified majority, the information on the voting behavior of the other delegation members was shown and the voting continued. So far the procedure follows SIM. In contrast to the former, the voting of the two delegations took place sequentially.[191] First, one delegation decided on the column (A, B or C). After this delegation had reached a decision, all six players were informed about the chosen column.[192] Then the other delegation voted on the row (1, 2 or 3). Again, the intersecting set determined the chosen alternative.

The two delegation procedures SIM and SEQ differed based on the degree of information players had available for their decision. Under SIM each delegation was not informed about the decision of the other delegation until it had reached an agreement of its own. Under SEQ the delegation which decided on the column "moves first" (this delegation is further referred to as first stage). The three participants of the first stage knew that the other delegation will be informed about their decision before they "move second" and vote on the row (this delegation is further referred to as second stage). As the second stage delegation already knew the results of the column vote, they faced a rather simple decision. For example, assume in the first stage column $\{B\}$ is selected. Then the second stage decision problem over the row (1, 2 or 3) is reduced to the alternative-specific question $\{B1\}$ vs. $\{B2\}$ vs. $\{B3\}$.

In a nutshell, my experiment focused on the separation of decision-making competences. I altered the voting procedure in terms of sequence and competences of the involved players (cf. Tab. 4.1). Under POL, my baseline, all six players decided simultaneously from nine alternatives. Each participant had one vote and could choose to allocate it to one specific alternative. SIM divided the group of six into two delegations of three players. Within each delegation votes were cast simultaneously, and both delegations held their ballots at the same time. The participants were now asked to choose a subset (either column or row) of alternatives. The intersection of the decisions of the two delegations

[191]The introduction of sequence was, by no means, only a small alternation. Game theory judges the timing of play as a fundamental characteristic of a game and classifies simultaneous and sequential games into different categories (cf. Dixit et al., 2009). In Chap. 5 I therefore introduce appropriately adapted game theoretic solution concepts for each design.

[192]The delegation voting second was informed about the subset chosen by the first delegation. Yet, they could only observe the collective decision but not which of the players cast their vote in favor or against the majority decision.

determined the final outcome. SEQ introduced sequence to the voting process. While within each delegation votes were still cast simultaneously, the voting of the two delegations took place sequentially. Most importantly, the second stage delegation knew the final decisions of the first stage delegation before holding its own ballot. Participants of the first stage delegation chose subsets while participants of the second stage delegation were asked to select a specific alternative out of the remaining subset (which always contained three alternatives).

Overall, changing the voting rules generated four distinct decision situations which differed with respect to their cognitive complexity and prevailing uncertainty: POL, SEQ first stage, SEQ second stage and SIM. Before I assign the respective manifestations of the two characteristics to these situations (cf. Tab. 4.2), I clearly distinguish the criteria from each other.[193]

COMPLEXITY

I start with complexity. This is no trivial task. Braumöller (2003, p. 212) attributed this to the fact that "the concept is so slippery." This holds even though highly complex and difficult-to-analyze dynamics are common in real-world processes (Lindgren, 1991).[194] Hinich (2008, p. 999) pointed out that "political and social games are so complex that the assumption of common knowledge that all actors know all the states of nature in the games and the conditional joint density of the states is grossly false." Accordingly, Koremenos (2008, p. 169) understood a complex problem "as uncertainty about behavior or the state of world, enforcement problem, or commitment problem." Another attempt to define complexity was made by Jervis (1997, p. 35) who referred to constellations where "the effect of one variable or characteristic can depend on which others are present." Yet, these perceptions are not unambiguous as they comprise a mixture of uncertainty, ignorance and interdependence as a cause. To structure my experimental design I strive rather to define complexity and uncertainty independently.

I agree with Braumöller (2003, p. 212) in that the most concrete and clear solution is provided by Ragin (1987) who assessed a (causal) complex case as one that "results from several different combinations of conditions" (Ragin, 1987, p. 20). He added further that "multiple causes interact with one another to produce effects, and the manner in which they interact is described by the logical operators 'and' and 'or"' (Ragin, 1987, p. 89-93).[195] In my experiment the decisions on column and row of a payoff table interacted to assemble the final outcome. The level of complexity resulted from the number of multiple factors which had to be taken into consideration. I did not try to measure complexity

[193]Cf. Oaksford and Chater (1998) for a broader overview and compilation of contributions on the cognitive science of human reasoning under uncertainty.

[194]Lindgren (1991) conducted a simulation model analysis which encounters several kinds of evolutionary phenomena. For example, these are periods of stasis, punctuated equilibria or large extinctions.

[195]Braumöller (2003, p. 212-213) pointed out that by using this conception many operational concepts as, e.g., substitutability or necessary and sufficient conditions can be understood as special cases of complexity.

in exact figures. Rather, I restricted my classification, when following this definition, to relative assessments.[196]

The splitting up of the decision-making competences implied a reduction in the number of options which participants could select (9 alternatives vs. 3 subsets). Yet, this was true only if supplemented by a sequence of decisions, as with simultaneous acting the other delegation does play a role. Under SIM the mutual dynamics kept up the level of complexity as nobody knew the result of the other delegation before reaching an agreement in its own. The tasks under POL and SIM were not identical, but each demanded sophisticated computations in its own way. Thus, the number of the interacting causes was reduced only under separated and sequenced decision-making, and so was the level of cognitive complexity. This is best illustrated when considering POL in comparison to the second stage under SEQ. Both situations allowed the participants to vote for a specific alternative. Yet, while the first comprised five other players and nine alternatives, the second contained two other players and three alternatives. The second stage under SEQ therefore just resembled a relatively simple copy of POL.

UNCERTAINTY

The second varying aspect in my design was the level of uncertainty.[197] Here, it is not difficult to find a definition; rather one must select the appropriate version out of the multitude of possibilities. Most important is the distinction between exogenous and endogenous uncertainty; i.e., between known and unknown probabilities for certain outcomes. This classification goes back to Knight (1921).[198] Under exogenous uncertainty the probabilities of occurrence for all possible events are given in advance. A lottery resembles such a state where chance and associated risk are common knowledge. "More generally, risky situations are games against nature" (Heinemann, 2005, p. 296). Contrary to this, under endogenous uncertainty the specific probabilities of an event are unknown. This arises, e.g., "in situations, where the outcome depends on social interaction" (Heinemann, 2005, p. 296).

In addition to these two clear-cut cases, Ellsberg (1961) associated situations which are characterized neither by known risks nor by complete uncertainty but by the lack of information about relative likelihoods with the term ambiguity. Here, Weber and Johnson (2008, p. 132) underlined clearly that "knowledge about the probability distribution of possible outcomes of a choice can lie anywhere on a continuum, from complete ignorance (not even the possible outcomes are known) at one end, through various degrees

[196] Cf. Stanovich and West (2000, p. 648-649) for a comprehensive discussion of statistical measures for cognitive ability.

[197] Cf. Weber and Johnson (2008) for a comprehensive assessment on decisions under uncertainty. The authors covered psychological, economic, and neuroeconomic explanations of risk preferences and also included an historical overview.

[198] Knight (1921) reached this conclusion when assessing business decisions which typically involve non-measurable risk, as they "deal with situations which are far too unique, generally speaking, for any sort of statistical tabulation to have any value for guidance. The conception of an objectively measurable probability or chance is simply inapplicable" (Knight, 1921, p. 231).

of partial ignorance (where outcomes may be known, but their probabilities not precisely specified, denoted as uncertainty or ambiguity), to risk (where the full outcome distribution is precisely specified), to certainty (where only a single, deterministic outcome is known to result)." In their comprehensive contribution on ambiguity Eichberger and Kelsey (2008) distinguished between the different terms in the following way: uncertainty is "a generic term to describe all states of information about probabilities. The term risk will be used when the relevant probabilities are known. Ambiguity will refer to situations where some or all of the relevant information about probabilities is missing. Choices are said to be ambiguous if they are influenced by events whose probabilities are unknown or difficult to determine" (Eichberger and Kelsey, 2008, p. 4).

Within the domain of uncertainty one must differentiate, for the reason of missing information as environmental, role or strategic uncertainty are possible. The first two can be excluded with respect to my experimental design, though for quite different reasons. Environmental uncertainty resembles a cornerstone in management decision-making (Holm et al., 2013, p. 2). Here, uncertainty arises from the fluctuating demand of customers (Bernstein and Federgruen, 2005), delays in project operations (Pich et al., 2002) or, more generally, unknown environmental conditions (e.g., Aldrich, 1979; Leblebici and Salancik, 1981).[199] For a laboratory investigation such sources of uncertainty are quite unlikely.[200]

Role uncertainty is "a commonly used experimental procedure. It consists of collecting from the same subject responses to tasks assigned to different roles, and letting a random mechanism determine which role's actions will be implemented" (Iriberri and Rey-Biel, 2011, p. 160). Yet, in my experiment subject were always aware of their role. Their identity (player number), the environment (decision-making procedure) and the earning possibilities (payoff table) were always known before they had to make a decision.[201] The controlling capabilities of a laboratory allowed me not only to prevent biasing influences but to set a specific common information level for all participants.

The term strategic uncertainty was first used by Van Huyck et al. (1991). Most importantly, the authors pointed out that "strategic uncertainty arises even in situations where objectives, feasible strategies, and institutions are completely specified and are common knowledge. [... Thus, it] should not be confused with uncertainty arising from incomplete information about other aspects of a decision maker's environment" (Van Huyck et al., 1991, p. 886).[202] While defining the basic principles, they did "not give a proper definition, but it seems clear that they mean the uncertainty arising from multiple equi-

[199] Morris and Shin (2002, p. 2) used also the term "'structural uncertainty" for uncertainty related to "the underlying fundamentals."

[200] This is especially true in comparison to potential causes within the experimental design.

[201] Before starting the experiment, a quiz tested the subjects' understanding of the upcoming game and secured their understanding of the displayed information.

[202] Van Huyck et al. (1991) referred prominently to (Sugden, 1989, p. 881) and his "lucid critique of the view that a rational decision maker can deduce a unique 'rational' strategy from the information contained in a complete information description of a game."

libria" (Heinemann et al., 2009, p. 181). This left room for interpretation considering a concrete definition.[203]

Holm et al. (2013, p. 3) described "competition (where the uncertainty concerns the individual's performance relative to others) and trust (where there is a 'social' risk that another party does not act favorably towards the trustee) to exemplify situations involving strategic forms of uncertainty." Thus, this type of uncertainty relates prominently to other players' actions (Andersson et al., 2012). Morris and Shin (2002, p. 2) extended its context to "uncertainty concerning the actions and beliefs (and beliefs about the beliefs) of others." Here, the multitude of equilibria arose out of players' own actions and their expectations concerning other players' actions. Many equilibria cause computational problems as the number of possible states that need to be considered is large. In such cases, "even if the underlying fundamentals of the problem were known for sure, the strategic uncertainty is still all-pervasive" (Morris and Shin, 2002, p. 2). Recent experimental studies have validated this influence of (strategic) uncertainty on behavior (e.g., Heinemann et al., 2009; Cabrales et al., 2010). More precisely, "in such situations, even the slightest uncertainty about other players' choices might lead a player to deviate from his or her equilibrium strategy" (Andersson et al., 2012, p. 1).[204] Even worse, a heterogeneous response to uncertainty provokes misunderstanding.[205] Of course, "public information reduces coordination failures and leads to more efficient strategies" (Heinemann et al., 2009, p. 184). Yet, often the provision of information is not possible.

The linkage between strategic uncertainty and ambiguity becomes clear when one considers their causation; both forms of uncertainty originate from the absence of relevant information (about probabilities). Interestingly, Heinemann et al. (2009, p. 2) found that a "subject's certainty equivalents of coordination games are positively related to certainty equivalents of lotteries[. ... Thus,] subjects who avoid risk or new experience also avoid strategic uncertainty. This suggests that subjects have similar perceptions of exogenous and strategic uncertainty if both situations are framed in a similar way."[206] By consid-

[203]Due to my experimental setting I could exclude concepts as "reverberant doubt" (cf. Hofstadter, 1985, p. 752-753), i.e., the fear of a non-realization, from my further discussion. As I always iterated ballots until a collective agreement is reached this specific form of strategic uncertainty is not relevant.

[204]Crawford and Haller (1990, p. 572) even stated that "this strategic uncertainty undermines the arguments that players should play according to any given equilibrium and even calls into question the rationale for playing an equilibrium strategy." Yet, this debate seems undecided as Andersson et al. (2012, p. 1) made the counter-argument that "in the laboratory, human subjects' behavior in games with multiple equilibria has also been found to be fairly stable and predictable in the aggregate".

[205]I restricted my analysis to ballots which yielded a collective result by the necessary majority of individual votes (cf. Sec. 4.3.1). This could not answer all questions as "the behavior of people in situations of risk and uncertainty is complex and multiply determined" (Weber and Johnson, 2008, p. 141). Further studies may look into the complete bargaining process and include non-decisive ballots. This would enable one to utilize the various tools provided by psychology and neuroeconomics as, e.g., one-step look-ahead algorithm (OSLA) or reinforcement learning (Chalkiadakis and Boutilier, 2004, 2008). Those instruments will then enable "a far more nuanced assessment and understanding of both general behavior patterns and individual or group differences in behavior" (Weber and Johnson, 2008, p. 141).

[206]In addition, Heinemann et al. (2009) investigated the "experience seeking" of their participants by using the Sensation Seeking Scale V (SSS-V, Zuckerman, 1994), a psychologist's measure to characterize personalities. They found that the scale resembled a good predictor for a subject's certainty equivalents.

ering strategic uncertainty in my experimental design I therefore also covered the wide range of exogenous uncertainty.

Following Brandenburger (1996) and Heinemann et al. (2009), I defined strategic uncertainty as uncertainty arising from missing information about the purposeful behavior of players in an interactive decision situation. In the words of Crawford and Haller (1990), players in a coordination experiment "may thus bear significant uncertainty about how other players will respond to its multiplicity of equilibria, even with complete information" (Crawford and Haller, 1990, p. 572). Therefore, the "coordination strategies reflect their uncertainty about how their partners will respond to multiple-equilibrium problems; this uncertainty constrains the statistical relationships between their strategy choices players can bring about" (Crawford and Haller, 1990, p. 571).

In my experiment, most of the voting record information was missing when subjects have to find a collective agreement under SIM. Here, each delegation was not informed about the decision of the other delegation until they reached an agreement themselves. Until the final choice was made, no player knew the specific consequences of their choice with certainty. This distinguished SIM from SEQ which applied also separated decision-making.[207] Here, the three participants of the first stage knew that the second stage delegation was informed about their choice before they had to decide themselves. Yet, while they were able to restrict the choices of their successors they were still left in the dark about the consequences of these actions. Thus, the first stage was associated with considerable uncertainty. Contrary, the second stage finalized the decision. Because the other delegation‘s vote was already known, the level of uncertainty was distinctly lower when players were part of the complete or, at least, the last step of the decision process. This holds true also for POL where no information was withheld from any player.

DECISION SITUATIONS

To summarize, POL was cognitively demanding because subjects had to consider five other players and nine alternatives. Yet, it did not resemble a nonseparable decision problem as the other two voting rules. Every subject was able to choose his preferred alternative unambiguously. Neither voting competences nor voting sequences required a separation of the decision into single issues.

Contrasting, SEQ had two stages. In the second stage subjects found themselves in a drastically simplified (i.e., less cognitive demanding) version of POL with only two other players and three alternatives. Subjects in the first stage had to anticipate the outcome of the second stage for each of the three possible subsets. Yet, at least they got a chance to set the agenda for the succeeding group; in the end, they knew which remaining subset their successors would vote on. This reduced the complexity in comparison to that faced by subjects under SIM, who had to consider mutual dynamics.

[207] The chosen form of separated decision-making excludes all forms of signaling through collusion (communication was forbidden) or reputation building (reshuffling after each round).

Under SIM subjects were also exposed to a high level of uncertainty about the behavior of others. They had to evaluate nine alternatives for six players in two delegations of three players acting at the same time without any further information. In the first stage under SEQ there was also no information on the behavior of its successors. Uncertainty is related to how far to the end of the decision-making a player was involved. This can be deduced from their choice opportunities. Could they only choose between subsets or were they allowed to select a specific alternative? To vote for a concrete alternative was only possible under POL or the second stage under SEQ. In relation to the other situations the level of uncertainty was rather low as in both cases the participants took part in the decision-making until the final and alternative-specific pick. Tab. 4.2 orders the decision situations based on cognitive complexity and uncertainty. The (relative) distinctions between the situations enabled me to identify the observed effects for the characteristics separately.

Table 4.2: Characteristics of the different decision situations

EXPLANATORY NOTE
The table points out the characteristics of the four different decision situations. They can be differentiated according to their cognitive complexity and prevailing uncertainty. The level of complexity refers to how demanding the computations were a subject had to perform. The level of uncertainty indicates to what extent a subject remained in the dark about the behavior of other participants.

		LEVEL OF UNCERTAINTY	
		lower	higher
LEVEL OF COMPLEXITY	lower	second stage sequential delegation	first stage sequential delegation
	higher	pooling	simultaneous delegation

Complexity and uncertainty are two main characteristics which come along with nonseparability. As discussed in Sec. 2.5, the previous (theoretical) literature attributed them to cause opportunities for strategic manipulation in form of a first-mover advantage (ARGUMENT 2) as well as sub-optimal outcomes (ARGUMENT 3). The first aim of my experiment was to examine these arguments empirically. In accordance with literature on laboratory experiments, the alterations in uncertainty and complexity led to further changes in subjects' behavior. Chap. 5 and Chap. 6 discuss this literature on behavioral patterns and their expected variation. Those will be formulated as specific expectations for my experiment. To investigate if these expectations are confirmed was the second aim of my empirical analysis; and thus, the theoretical discussion about nonseparability is supplemented with further arguments on its impact.

4.2.7 Two procedures per session

I implemented two procedures in every experimental session. At the beginning of the experiment, the subjects were informed that the experiment would have one specific decision rule for some rounds and that this rule would be altered in later rounds. They were not told in advance how many rounds they would play under either rule. In fact, the number of rounds for the two procedures was not even known to the instructor. Instead, they were endogenously defined while the experiment was running.

Before conducting a session, the laboratory had to be booked for the corresponding time slot. In this duration, the participants were brought into the lab, the instructions were explained, several rounds of the experiment were played, the participants filled out a questionnaire and all subjects received their payoffs. The available time slot for the voting experiment was approximately $\frac{2}{3}$ of the overall reserved time. The procedure changed after half of this time had gone by. The number of rounds played so far was not decisive. If the committees needed more time to reach a decision they played fewer rounds. If they agreed quickly they participated in a higher number of rounds.

As I implemented two decision rules in every session, the examination of the data was a mix of *between-subject* and *within-subject* analysis[208]. In a within-subjects design, each participant is tested under each procedure (Morton and Williams, 2010, p. 86-87). The results of the same subject under the different procedures are compared. In a between-subjects design, each participant is tested under one procedure only (ibid.). The results of different subjects under each condition are compared. My experiment resembled neither exactly. No subject is tested under all or only one procedure. This followed the idea of randomized partial counterbalancing (cf. Mitchell, 2003, Chap. 13, p. 485).

In my analysis of the experimental data I used both methods as they complement each other.[209] The between-subject design implies fewer biases like EDE, treatment confounding, etc. The within-subject design reduces individual variability and noise. This enabled also to obtain a within estimate of the treatment effect.[210] "The accuracy of this approach depends on whether any order biases cancel one another out across the two orders" (Charness et al., 2012, p. 6). A requirement for combining both methods is that subjects are tested under different modifications of the same factor (MacKenzie, 2012). In my experiment the treatment intervention altered the decision-making procedure.[211] For the between-subject analysis I used only the data gathered under the first decision rule in each session. For the within-subject analysis I utilized also the data gathered under the second decision rule compared to the first procedure in each session.[212]

4.2.8 Payment mechanism

It is still an ongoing discussion whether and how to pay participants in experiments. Amir et al. (2012) investigated if financial motivation matters at all. Therefore, they re-

[208] Cf. Martin (2008, Chap. 8, p. 148-170) for an excellent juxtaposition of the two approaches.

[209] Cf. Charness et al. (2012) for a discussion of the advantages and disadvantages of the methods.

[210] Isaac and Walker (1988a,b) analyzed how individuals change their behavior in response to a change in stakes in public good experiments.

[211] MacKenzie (2012) plainly illustrated the necessary conditions for a combined analysis: "One group of participants is tested under condition A, a separate group is tested under condition B, and so on. The assumption here is that condition A and condition B are different levels of the same factor. For example the factor might be device and the levels might be mouse, trackball, and touchpad. In experiments with more than one factor, it is possible to use a within-subjects (repeated-measures) assignment for the levels of one factor and a between-subjects assignment for the levels of another factor."

[212] The distinction between first and second decision rule also accounts for possible learning or experience effects of having played a different procedure before.

peated canonical economic games online using very low stakes.[213] They found that their results were comparable to those run in laboratory settings. Camerer and Hogarth (1999) argue that no replication of an experimental study has achieved more theory conformation or less rationality violations by purely raising incentives. The authors do not call for more studies without financial incentives generally. The benchmark is always set by "previous [related] research showing that financial incentives did not matter in their task" (Camerer and Hogarth, 1999, p. 25). As I link my results to previous work on collective decision (cf. Sec. 4.1.2) I followed their lead by using financial incentives.

If participants would be paid for every round using an accumulated payoff mechanism (APM), the experiment would clearly run into wealth, hedging or portfolio effects (cf. Cox, 2010; Ham et al., 2005; Heinemann, 2005). I would also not be able to regard every round as a single game because of spill-over effects between decision problems. In order to avoid such effects, the total earnings were determined by a random round payoff mechanism (RRPM)[214]. Using this mechanism one or a few rounds are randomly chosen as the basis for a subjects payment (Morton and Williams, 2010, p. 279).

In addition to preventing wealth effects RRPM has the additional advantages that monetary payoffs in each period can be raised. This may increase the salience subjects assign to their choices (Morton and Williams, 2010, p. 279). So the payment mechanism influences a subject's perception: every round comes closest to a single one-shot game. This follows Kuziemko et al. (2011, p. 17) in abstaining from dynamic payoff accumulations over rounds for "making the experiment a series of one-shot games". The advantage of such a series is that it is not necessary "to make assumptions regarding players time horizons" (ibid.). This facilitates the analysis as it is not easy to distinguish what determines subjects' time horizon or how they try to maximize their payoffs. Benartzi and Thaler (1995) argued that individuals just always maximize their current payoffs. This holds true even in experiments where the defining outcome is the future final round. Gneezy and Potters (1997) showed that individuals maximize a specific "evaluation period". In my experiment subjects neither knew the end of the experiment nor which rounds contributed to their payoff. Each round should therefore have constituted such an evaluation period and have been played as if it was one-shot. This is in line with Camerer et al. (1993, p. 44) who concluded that even in a simple sequential game "subjects concentrated on the current round when making decisions." In summary, experimental subjects tend to maximize current payoffs even in a setting over multiple rounds or when the received payoff is explicitly based on the final balance.[215]

[213] More precisely, Amir et al. (2012) implemented two payoff conditions: i) a "stakes condition", in which a subject's payoff was based on the outcome of the game (max. $1), and ii) and a "no-stakes condition", in which a subjects payoff was unaffected by the outcome of the game. Overall, they found some but little difference in behavior between the stakes and no-stakes conditions.

[214] This way of paying subjects is also called random lottery incentive mechanism (Grether and Plott, 1979).

[215] Kuziemko et al. (2011) explained this behavior with uncertainty about the experimental setting. Subjects might value the current payoffs more because their distribution is known with certainty while future rounds are undetermined.

Arzieli et al. (2012) explained that repeated experiments comprehend two relevant elements which determine the choices participants make. The first is the decision problem, and the second the payment mechanism. These aspects cannot be separated from one another, as the implication of a specific design is changed by different payment schemes (Chandrasekhar and Xandri, 2012). Thus, the same experiment would yield different results if it is played using either an APM or RRPM (Lee, 2008).[216] An experiment therefore needs an "incentive compatible payment mechanism" (Arzieli et al., 2012, p. 15). A subject's overall optimal strategy also has to be their optimal choice when looking at each single decision problem individually. This follows the logic of subgame-perfect equilibria (SPE) (Stahl and Haruvy, 2008). Looking at different possible payment mechanisms Sherstyuk et al. (2011) conclude that only the payment for one randomly-selected choice satisfies this condition without any further assumptions.

At least, similar approaches have been pursued in previous work on collective decision-making (cf. Sec. 4.1.2). Here, S&K (2010) paid for 3 randomly selected decisions out of 20 rounds. In the experiment conducted by Charness and Rabin (2002) subjects made between two and eight choices and were paid for one or two of their choices, which were selected at random. Thus, it was reasonable for my experiment to use an RRPM and to pay for a subset of all decisions made. This left open the question of how many rounds the subset should consist. In the end, the total earnings were determined by three randomly selected rounds. The points earned in these rounds were paid in cash with an exchange rate of 5 Points to € 1. Additionally, all participants received a "show-up fee" of € 4, which was also paid to all over-recruited subjects (following Morton and Williams, 2010, Sec. 10.2.3, p. 282).[217] The motivation to include the fee was mainly that it ensured that all participants received at least some earnings. Also, current research by Azar (2010, p. 24) suggested that such a fixed payment "does not affect the magnitude of the pay-for-performance component".[218] Thus, a fixed show-up fee did not diminish the induced incentives. I chose the exchange rate (and show-up fee) in a way that the average compensation per subject and hour should correspond to the average hourly wage of a scientific assistant.[219] All these payoff regulations were common knowledge to the participants and stated at the beginning of each session.

The decision to pay the earnings from three random rounds, instead of just one, was made in response to the subjects' reactions after the pilot session. The payoff tables of my experiment displayed a wide range of possible payoff distributions (Sec. A.8). When

[216] Lee (2008) directly compared both mechanisms and found evidence that the payment mechanism influenced subject's choices. With APM he detected wealth effects which were not observed using RRPM.

[217] For every session with 18 subjects I invited originally 21 subjects. Thanks to this reserve only one scheduled session could not take place.

[218] Azar (2010, p. 1) defined "relative thinking" as to consider relative differences although only absolute differences are relevant. Using laboratory experiments he determined that this kind of thinking is limited in scope (e.g., when people compare prices of goods) and does not apply when considering payments for task performances.

[219] This followed the AWI (2013) rules and conditions of participating in experiments and constitutes a widely used standard procedure in conducting an experiment (e.g., CESS, 2012).

they were paid for one round, the subjects complained about the "uneven" or "unfair" distribution of points and the "lack of opportunities" on a good performance. To them it seemed that the draw of the one payment round was the only important factor and the act of voting was more just an accessory. This can hardly account as an incentive-compatible payment mechanism for a voting experiment. The payment of three rounds was chosen to ensure a high motivation of the subjects during the ballots.

This RRPM or "random multiple decision selection mechanisms" (RMDSM, Arzieli et al., 2012, p. 20) presupposes the additional assumption that "no complementarities exist between the different rounds" (Arzieli et al., 2012, p. 20). In my experimental context this means that, for example, a subject who prefers alternative $\{A1\}$ to alternative $\{B2\}$ in round 3 ($R_3 : A1_3 \succeq B2_3$) and alternative $\{A3\}$ to alternative $C3$ in round 5 ($R_5 : A3_5 \succeq C3_5$), will prefer $\{A1\}$ in round 3 *and* $\{A3\}$ in round 5 $(A1_3; A3_5)$ to (i) $\{B2\}$ in round 3 *and* $\{A3\}$ in round 5 $(B2_3; A3_5)$ and (ii) $\{A1\}$ in round 3 *and* $\{C3\}$ in round 5 $(A1_3; C3_5)$ and (iii) to $\{B2\}$ in round 3 *and* $\{C3\}$ in round 5 $(B2_3; C3)$. Thus, preferences had to be independent of the existence and results of other rounds. This had to account to all possible combination of rounds and preferred alternatives.

4.2.9 The subjects

The subjects were recruited using the ORSEE software out of a subject pool of the laboratory run by the Department of Economics at the University of Heidelberg. The pool contained about 1,300 people. Nearly all of them were either students at the Universities of Heidelberg or Mannheim. The subject pool is advertised by flyer distribution as well as information sessions in first year bachelor classes. Those sessions inform about the experimental laboratory and the financial possibilities. Subjects who are interested register themselves online in the ORSEE data base[220]. Of course, this leaves the subject sample with the same kind of (expected) selection bias most other pools have. Furthermore, the self-registration of the subjects corresponds perfectly to the case of a "volunteer". Rosenthal and Rosnow (2009, Chap. 3) showed that behavioral research draws their samples in general from such populations of volunteers. Much more important than just their self-motivation, those individuals might differ from those not finding their way into research.

However, current research suggests that experimental subjects are an appropriate subject pool. They are not at all significantly different from the general population of which they were drawn (Branas-Garza et al., 2012). Thus, my findings for the self-selected student sample are generalizable to the student population; but can the results be generalized beyond? Do, e.g., professionals behave differently than students? Here, Frechette (2011) reviewed experiments that included both types of subjects and found, on a general level, that results of both subject pools lead to similar conclusions.[221]

[220] Registration page: `http://147.142.190.243/orsee/public`.

[221] Frechette (2011, p. 1) defined professionals "as people working in an industry where the game under study is thought to be relevant."

Cleave et al. (2010) investigated if social and risk preferences of participants in laboratory experiments represent the preferences of the population from which they are recruited. Using 1,173 students they found that the preferences of the participants are not significantly different from the preferences of the population they were recruited from. It seems that the statement of selection bias is not always valid. Of course, Cleave et al. (2010, p. 1) limited their conclusions as they "fail to find selection bias based on social and risk preferences." This tells us nothing about other characteristics. To exclude a selection bias the specific patterns of the research question and the sample have to be considered.

Social preferences in general have received considerable attention among economists in recent years. Falk et al. (2011) pointed out that this research was nearly exclusively focused on student samples. The authors conducted two studies to find out if this distorted the results. They found "that self-selection does not significantly bias the social preferences measured in the laboratory" (Falk et al., 2011, p. 13) and "that student participants and non-student subjects show very similar behavioral patterns" (ibid.). In fact, non-students showed significantly more social behavior. The authors suggested therefore that results from student samples might be seen as a lower bound for the importance of pro-social behavior. Therefore, I do not regard the high percentage of students in the subject pool as a problem when it comes to social preferences.

These findings are relatively new and the broad use of student samples was due to practical considerations. Druckman and Kam (2009) offered an excellent discussion of this "convenience sampling"[222]. In the lab this refers to the aspects of proximity, availability and affordability. Henrich et al. (2010) conducted a meta-analysis on numerous psychological studies and found that 96% of all test subjects were from western industrialized countries, which account only for 12% of the world population.[223] The authors referred to this group as "WEIRD" which stands for "Western, Educated, Industrialized, Rich and Democratic". Such divergence is a problem if the group of subjects behaves differently than the actual target population of the experimental research. Druckman and Kam (2009) argued that this depends on the underlying data generating process. If "the treatment effect is the same across populations, the nature of a particular sample is largely irrelevant for establishing that effect" (Druckman and Kam, 2009, p. 12).[224]

Following Levitt and List (2007), I argue that whether behavior inside the laboratory is a good indicator of behavior outside the laboratory depends on the nature of the decision environment. The authors emphasized in particular the extent of scrutiny by others. Therefore, laboratory experiments may "prove to be better suited for naturally-occurring

[222]"Convenience sampling is a non-probability sampling technique where subjects are selected because of their convenient accessibility and proximity to the researcher" (Castillo, 2009).

[223]A second finding in Henrich et al. (2010) was that 67% of the subjects were themselves students of psychology at American universities.

[224]Druckman and Kam (2009, p. 12) clarified further that "if the underlying data generating process is characterized by a homogeneous treatment effect [...], then any convenience sample should produce an unbiased estimate of that single treatment effect, and, thus, the results from any convenience sample should be easily generalizable to any other group of individuals."

settings in which there is a high degree of scrutiny of actions (e.g., employer-employee relationships, family interactions), or an emphasis on process (e.g., politics, judicial proceedings)" (Levitt and List, 2006, p. 4). Politics is closely monitored by the media[225] and the public considers its structure and process very thoroughly (e.g., McCubbins et al., 1989; Lane and Ersson, 2000).[226] Thus, out of the variety of real-world circumstances the decision-making in politics may be one of the most suited processes for laboratory scrutiny. The emphasis on process corresponds also well with my treatment, the modification of the decision-making process for a group of individuals.[227]

To be honest, it would have been nearly impossible to bypass the problems of self-selection or convenience sampling. Either way would have called for a considerable amount of money and time. In Sec. 9.2 I discuss the possibility of further research to include online experiments or "experimental turks"[228] (Rand, 2012) into the subject pool. This would broaden the pool enormously, but it would rather not solve the potential problem of selection bias.

4.3 Implications for the empirical analysis

Sec. 4.2 provides a comprehensive overview of my experimental design. It becomes clear that it was no ordinary lab experiment as it followed the concept of "randomization within constraints" (Martin, 2008, p. 30ff). More precisely, I established some constraints (e.g., up to 18 subjects and two procedures per session in a fixed period of time) and made random assignments within these constraints (e.g., random chosen payoff table, group composition and number of rounds). The randomization enabled generalizability and excluded confounding factors from the analysis (Bernauer et al., 2009, p. 91) while the control secured an unbiased database.[229]

The augmenting features of my design allowed me to look into questions so far not investigated by other experimental studies. I compare the effect of several decision-making structures and contrast separable and nonseparable decision problems directly. The comparison is performed on both the collective and the individual level. Also, I do not limit

[225] Druckman and Parkin (2005) argued that the media does not only monitor but influence politics.

[226] Sen (1997, p. 745) analyzed maximizing behavior and identified "process significance" as one of the most relevant aspects in the act of choosing.

[227] Overall, Levitt and List (2007, p. 154) examined five factors and their influence on decision-making: ethical considerations, scrutiny of one's actions, experimental context, self-selection of participants and stakes of the game. Looking at their results the authors concluded that "being monitored proves to be the critical factor influencing behavior in this study" (Levitt and List, 2007, p. 160).

[228] Mason and Suri (2012) discussed the use of Amazon's Mechanical Turk in detail. The central purpose of their paper is to demonstrate the usefulness of this approach for behavioral research. In short, the online labor market enables employers to post and workers to choose jobs. Thus, the platform provides a large, stable, and diverse subject pool. They judged the markets to offer researchers at low cost a fast iteration between developing theory and executing experiments. This lowers the barrier of entry for researchers. Yet, there are also problems with respect to monitoring, simultaneous participation, data privacy, etc.

[229] In particular the randomized assignment to the groups was essential as it enabled to control for the individual characteristics of the subjects (cf. Morton and Williams, 2012, p. 20).

myself to one specific constellation, but use several problem specifications (i.e., characteristics of the payoff tables). Yet, as mentioned already, these features also come with some drawbacks. The specific experimental design has to be taken into account when analyzing the resulting data. The following sections discuss the implications and corresponding adjustments of the analysis.

4.3.1 Individual and collective level

While committee voting can be viewed as a standard game for a laboratory experiment,[230] the split-up decision mechanism added a new pattern. There is experimental research on collective decision-making but the overwhelming majority of laboratory work is focused on individual behavior. My setting enabled the collection of collective and individual data in the same experiment. Contrasting them follows the idea of Halfpenny and Taylor (1973) and will bring new insights.

The collective level is important as the decision-making had group-wide (which means in the realm of politics commonly society-wide) consequences for the allocation and distribution of wealth. Delegation is, in general, expected to increase efficiency and stability (Bendor and Meirowitz, 2004). However, it might as well affect effectiveness and equity. This possible trade-off complicates the assessment of the appropriate decision-making procedure. The individual level allowed me to look through the noise of a collective majority decision. Predictions of individual behavior are far more reliable than forecast based on collective data where a minority of "renegade individuals [...] can upset these predictions rather dramatically" (Halfpenny and Taylor, 1973, p. 444). As all procedures used majority rule, not all participants had to agree on one single alternative. The threshold for reaching a decision was always a qualified majority of $\frac{2}{3}$ of all votes.[231] Studies which only consider the collective level would miss the votes of the outvoted subjects.

This juxtaposition of collective and individual level had an impact on the empirical analysis. I had to restrict my individual data to the final ballot in order to also have settled collective result as counterpart for the computations. As subjects voted in every round until a collective decision was reached, this concerns an individual which either belongs to the majority that reached the threshold or one that gets outvoted.

As mentioned before, laboratory experiments facilitate the monitoring of every step of decision-making (Ordeshook and Winer, 1980, p. 730). Thus, in my experiment I recorded

[230] The variety of games used in experimental laboratory research is overwhelming. Frequently used experimental settings include market organizations, individual decision-making, bargaining behavior, auction markets, coordination games, committee (or group) voting, public good games, dictator games, ultimatum games, trust games, etc. (for an overview cf. Plott and Smith, 2008, p. xii). This compound can be used to investigate all types of set-ups like games which are cooperative or non-cooperative, iterated or one-shot, finitely or infinitely repeated, symmetric or asymmetric, zero-sum (i.e., constant-sum) or non-zero-sum, simultaneous or sequential, played with perfect or imperfect information, played by a single player or by multiple players, etc.

[231] This was the lowest common denominator of all procedures: four out of six votes under POL and two out of three votes in both delegation procedures.

all individual votes which formed the collective decisions as well as all ballots which did not lead to a collective agreement and forced a new vote. To investigate the unification process is an exciting research question in itself. If and how players use their votes as a signal? What beliefs about other players do they have and how are those updated? Following this idea, Sec. 9.2.3 discusses in detail the possibilities of communication between the participants of laboratory experiments. I particularly consider the issues of recording and evaluation. However, this is material and subject for future research. The current study focuses on the last ballot and its collective as well as individual decisions.

4.3.2 Statistical independence of observations

Every experimental session lasted for multiple experimental rounds. In between the rounds the subjects were randomly reshuffled between groups which had to reach a decision collectively. Even though reputation building in the course of the rounds of the experiment was not possible (cf. Sec. 4.2.2), the single rounds of a session are not statistically independent from each other. Measuring a subject repeatedly (i.e., over the single rounds) leads to a non-independence of observations, the so called "session-effect" (Frechette, 2012).

Frechette (2012, p. 485) nicely illustrated this effect with the analogy of participants as "multiple members of a family [where] observations from siblings might exhibit more correlation than those from individuals across households." More precisely, the "session-effect problem is defined as a within session correlation in the variable of interest (or the residual) once the relevant factors are controlled for" (Frechette, 2012, p. 485). The dependence of observations has to be taken into account, regardless of whether the analysis uses parametric or nonparametric tests (Siegel, 2001).

Statistically independent observations can only be obtained at the session level. This leaves two possible solutions to investigate the results at both the collective and the individual level. Firstly, regression analysis enables the use of pooled data from all sessions of a given treatment by considering clustering at the subject or session level. Secondly, (non)parametric tests can be based not on single rounds but on session averages per treatment of the variables in question;[232] even though this has the disadvantage that it considerably reduces the number of observations (i.e., from the number of rounds played to the number of sessions).[233]

4.3.3 Rather qualitative than quantitative findings

It is always difficult to link experimental findings to the real world. Morton and Williams (2010, p. 196) concluded that "the proof of external validity is always empirical". Thus,

[232] Frechette (2012) discussed in detail the sources and implications of session-effects. In particular, he also pointed out that using session averages may not solve the dependence problem.

[233] Cf. Vanberg (2008) for an application of the two approaches.

further experiments with variations of, e.g., the target population, subject recruiting and experimental method are essential to verify the insights found (cf. Shadish et al., 2002, p. 21). For example, in the case of social preferences the work of Carpenter and Seki (2005) supported the external validity with a field experiment. However, the authors also admitted that the exact laboratory measurements seem to have less in common with the data gathered in "daily work lives" (Carpenter and Seki, 2005, p. 20).

Levitt and List (2006, 2007) strongly emphasized the orientation of laboratory experiments on qualitative insights. Taking into account the findings of neighboring disciplines they pointed out that "the wealth of psychological literature suggests that there is only weak evidence of cross-situational consistency of behavior" (Levitt and List, 2007, p. 160). This insight is strongly linked to the aim of the experiment; "if the role of experiments shifts from testing theories to motivating the development of new theories [this] has the danger of creating its own world" (Schram, 2005, p. 236) which prohibits the generalizability of findings.

Quite apart from reality, already changing experimental environments resulted in varying behavior and different parameter measurements. Charness et al. (2007a) conducted "lost-wallet game" experiments in the laboratory and through the internet. The difference between the settings is that "the internet methodology increases the social distance to a high degree" (Charness et al., 2007a, p. 101). In the laboratory the participants observed who else took part in the experiment, i.e., they recognized that all co-players are students[234]. But in experiments conducted through the web the other participants could be anybody. The authors found that differences in behavior are linked to the variation of the social distance.[235]

My analysis and conclusions were therefore not interested in "quantitative magnitudes" (Levitt and List, 2006, p. 5), i.e., reporting exact estimates for behavioral parameters. Rather, my goal was to identify and explain "qualitative findings" (Levitt and List, 2006, p. 5), i.e., general patterns of behavior. I focused on differences between my treatments which were all accomplished in the laboratory. Thus, not the absolute but relative values are decisive, as those are much more credible and generalizable (Levitt and List, 2007).

4.3.4 Responsibility for the results

SIM was a special case when it comes to the question of responsibility for the collective outcome. Here, no specific alternative was selected by any subject; a unique feature of this procedure. Instead, the two delegations each chose a set whose intersection then

[234]For a study on the effect of performing in an experiment together with classmates (though not at the university, but in school) cf. Belot and Van De Ven (2009).

[235]Research on social distance represents a wide field. Some contributions referred simply to the impact of known family names of co-players (Charness and Gneezy, 2008). Others looked worldwide for comparisons. For example, Ruffle and Sosis (2006) compared Israeli kibbutz members and city residents while Bouckaert and Dhaene (2004) focused on inter-ethnic aspects between Belgian and Turkish businessmen.

determined the final result. No delegation was informed about the vote of the other delegation until the own collective decision was reached.[236]

During the experiment always an equal number of column and row delegations voted simultaneously. Their belonging to a corresponding delegation was set in advance by the matching algorithm. I used the resulting outcome of the intersection to determine the earned points of the subjects for that round. Yet, the correspondence of two delegations was set by chance. For example, when a column delegation voted for set $\{A\}$ and the matching procedure assigned a row delegation that voted for set $\{3\}$, then alternative $\{A3\}$ was the collective outcome for the whole group, i.e., both delegations. But when the column delegation would have been matched with a row delegation that voted for set $\{1\}$ the result of the voting procedure would be alternative $\{A1\}$. Thus, the empirical analysis had to take the matching procedure into account.

I implemented this by using for each alternative not its observed frequency of occurrence but its conditional outcome probability. This combined the choices of each column and each row delegation that voted on a payoff table in any of the 13 experimental sessions. Thus, I calculated for both delegations over all rounds under SIM for each payoff table the probabilities for choosing one of the three sets. For example, the probability of alternative $\{B2\}$ to be the selected as final outcome is the probability of set $\{B\}$ being selected by the column delegation times the probability of set $\{2\}$ being the outcome selected by the row delegation.[237] When I looked into the decision made under the different procedures I therefore used this probabilistic result for SIM. This means, e.g., that the points obtained through a group from a payoff table correspond to the respective selection probability weighted average over all alternatives of a payoff table.

4.3.5 Summary of key characteristics

This section summarizes the discussed key characteristics of the experimental design and their implications for the empirical analysis. Tab. 4.3 lists the characteristics one by one and states the appropriate empirical response.

The empirical analysis in the next two chapters follows these requirements. First, I take a look at the collective level and the influence of institutional rules on collective decisions in Chap. 5. Then, I alter the focus of the analysis and look for the determinants of individual choices in Chap. 6. Previously, Sec. 4.4 provides a descriptive overview of the collected observations as starter.

[236] Of course, subjects may have inferred what the other subjects were about to vote, but the actual result of the other group was unknown until a delegation's own decision was finalized. Chap. 6 looks more into the matter of sophisticated behavior, i.e., anticipatory voting.

[237] For example, assume that a payoff table was used 13 times under SIM. Of the 13 column delegations seven choose set$\{B\}$ and of the 13 row delegations four choose set$\{2\}$. Hence, the probability of observing alternative$\{B2\}$ as final outcome equals $\frac{7}{13} \times \frac{4}{13} = \frac{28}{169} = 16.6\%$.

Table 4.3: Design characteristics and their impact on the empirical analysis

EXPLANATORY NOTE
The table lists specific details of my experimental design and how the empirical data analysis took them into account. All characteristics have been discussed in earlier sections of this chapter which are indicated.

DESIGN CHARACTERISTIC	IMPLICATION FOR THE EMPIRICAL DATA ANALYSIS
multiple payoff tables and their random assignment (Sec. 4.2.4)	requires a normalization of the payoff tables
multiple rounds (Sec. 4.2.5)	requires a trend analysis of the results
split of voting competence (Sec. 4.2.6)	requires a distinction between delegation and group level
voting sequence (Sec. 4.2.6)	requires a distinction between first stage and second stage under sequential delegation
two decision rules per session (Sec. 4.2.7)	requires a within-subject and between-subject comparison
contrasting individual and collective level (Sec. 4.3.1)	restricts the individual data to the last ballot in each round
no statistical independence across rounds (Sec. 4.3.2)	restricts (non)parametric tests to session averages
social distance within the laboratory (Sec. 4.3.3)	restricts the findings to relative differences among the various procedures
simultaneous delegation used the intersection of sets as collective outcome (Sec. 4.3.4)	requires the usage of conditional outcome probabilities under simultaneous delegation

4.4 Descriptive information

Overall, I conducted 13 experimental sessions between May 2011 and February 2012 with 168 participants. Subjects were 57% female and on average 22.8 years old. Nearly 88% were 25 or younger. Of all subjects 92% were students, of which 46% studied economics. The mean semester of all student subjects was 4.5 (SD 3.3). On average, subjects had participated in 4.7 (SD 3.5) experiments before.[238] The subjects played between 5 and 16 rounds which took between 1:10h and 2:45h.[239] Average payoffs for a subject equaled 8.97€/h (SD 1.76€/h). Including the expenditures for a pilot session in December 2010 the expenses for the experiment sum up to € 3.422.

The distribution of voting rules in my experimental sessions is shown in Tab. 4.4. In every session I applied two different rules. In 5 of the 13 sessions I used POL and SIM as well as POL and SEQ. In the remaining 3 sessions the subjects played under SIM and SEQ.

Over all sessions I obtained data on 221 group decision, of them I collected 76 under POL, 85 under SIM and 60 under SEQ. This represents the number of collective results, but I have distinctly more individual data. Every group decision consists of six individual votes which leaves me with 1326 individual votes that constituted these collective

[238] The demographic information was obtained as part of a post-experiment survey. I discuss its results in detail in Chap. 7.

[239] This time period extends from the entry of the first participant into the laboratory to the payment of the last participant. It includes in addition to the actual voting experiment the assignment of cubicals, the handing out of instructions, answering questions, etc. The variation resulted from several reasons such as, e.g., how long it took the participants to answer the post-experiment survey or how many comments were made by the subjects during the payout phase (cf. Sec. 7.1).

results.[240] As discussed before (cf. Sec. 4.3.2), these are not statistically independent observations which can only be obtained at the session level. Here, the mean values were calculated based on average on 8.50 (SD 3.20) group decisions.[241]

Table 4.4: Data structure at the session level

EXPLANATORY NOTE
The table depicts the voting rules used on the individual sessions of the experiment. Under each rule the session averages of the variables in question can be calculated to secure statistically independent observations.

		NUMBER OF EXPERIMENTAL SESSION													
		1	2	3	4	5	6	7	8	9	10	11	12	13	Σ
PROCEDURE	pooling	X	X	X	X	X				X	X	X	X	X	10
	simultaneous delegation	X	X	X	X	X	X	X	X						8
	sequential delegation						X	X	X	X	X	X	X	X	8

Next, I look more thoroughly into the data structure of the observations. Tab. A.7 shows the frequency of use of each of the payoff tables. The random assignment rule for the payoff tables resulted in an uneven distribution of the tables across the procedures. More interesting than the usage of single tables under each voting rule is the distribution of the characteristics of the payoff tables, i.e., if the tables are (non-)constant-sum games and if they contain a core alternative (cf. Sec. 4.2.4). Yet, for the discussion of the core concept under all three procedures additional derivations are necessary. I explain the required considerations and describe its application in detail in Sec. 5.2. For now, I restrict the overview to the criterion of constant-sum or non-constant-sum game and, what is also important, if a procedure was used as first or second decision rule within an experimental session (cf. Sec. 4.2.7).

Tab. 4.5 gives an overview of the corresponding data structure for between-subject and within-subject analysis. The between-subject analysis used the data gathered under the first decision rule in each session; i.e., 28 POL, 60 SIM and 38 SEQ decisions.

Table 4.5: Data structure according to the order of decision rules

EXPLANATORY NOTE
The table depicts the number of observations in terms of collective decisions made in the experiment. The observations are divided according to procedure and first or second decision rule of a session.

		DECISION RULE		
		first	second	Σ
PROCEDURE	pooling	28	48	76
	simultaneous delegation	60	25	85
	sequential delegation	38	22	60
	Σ	126	95	221

For the within-subject analysis I could also use the observations gathered under the second decision rule of a session. Here, my data comprised 48 POL, 25 SIM and 22 SEQ results. The juxtaposition of POL with SIM or SEQ enabled me to investigate how subjects react to the divided decision-making (competencies or sequence). When conducting POL after SIM, the divided competencies vanished. When POL followed after SEQ, both

[240] In addition, when a ballot did not reach the necessary majority another ballot was held. Thus, I recorded overall 4695 individual votes.

[241] Looking at each voting rule separately, the session level data comprised on average 7.60 (SD 3.88) POL, 10.63 (SD 2.69) SIM and 7.50 (SD 0.87) SEQ collective results.

divided competencies and sequential voting were removed. Thus, of 48 POL results, 24 took place after SIM and 24 after SEQ. Reversing the sequence, I conducted 25 SIM rounds after POL to investigate what patterns of voting behavior change when decision competencies are split up. Furthermore, 22 SEQ rounds were run after SIM to analyze what happens if I keep divided competencies and split up the decision sequence.

As explained in Sec. 4.2.4, the experiment contained constant-sum as well as non-constant-sum payoff tables.[242] Tab. 4.6 displays the frequency of the experimental observations separated by this criterion (following Zizzo and Tan, 2009). The distribution of constant-sum and non-constant-sum tables shows nearly equal frequencies for the different procedures. 27 of 76 POL (35.5%), 35 of 85 SIM (41.2%) and 22 of 60 SEQ (36.6%) results emerged from rounds which used constant-sum tables.

Table 4.6: Data structure according to constant-sum and non-constant-sum table

EXPLANATORY NOTE
The table depicts the number of observations in terms of collective decisions made in the experiment. The observations are divided according to procedure and constant-sum or non-constant-sum payoff table.

		CONSTANT-SUM TABLE		
		yes	no	Σ
PROCEDURE	pooling	27	49	76
	simultaneous delegation	35	50	85
	sequential delegation	22	38	60
	Σ	84	137	221

Of course, these different payoff table characteristics can be combined. Tab. 4.7 separates the number of observations according to whether it is the first or second decision rule of the session and whether a constant-sum and non-constant-sum table is used. Taking both aspects together, e.g., 12 POL, 26 SIM and 12 SEQ results were obtained from a constant-sum table used under the first decision rule. The fluctuation in the number of observations resulted from the random assignment of payoff tables (cf. Sec. A.11).

Table 4.7: Data structure according to the order of decision rules and constant-sum or non-constant-sum table

EXPLANATORY NOTE
The table depicts the number of observations in terms of collective decisions made in the experiment. The observations are divided according to procedure, first or second decision rule of a session and constant-sum or non-constant-sum payoff table.

		CONSTANT-SUM TABLE						
		yes			no			Σ
		DECISION RULE			DECISION RULE			
		first	second	Σ	first	second	Σ	
PROCEDURE	pooling	12	15	27	16	33	49	76
	simultaneous delegation	26	9	35	34	16	50	85
	sequential delegation	12	10	22	26	12	38	60
	Σ	50	34	84	76	61	137	221

To secure a valid empirical analysis I list in the following, always explicitly, the respective number of observations. This refers in Chap. 5 to the amount of collective decisions and in Chap. 6 to the number of individual votes.

[242] In a constant-sum tables the overall sum of each alternative is the same; in non-constant-sum tables the overall sum varies between alternatives (cf. Sec. 4.2.4).

5 The influence of institutional rules on collective decisions

In this chapter I take a look at the decisions made in the experiment at the collective level. This addresses the first part of my second key question, which is whether and, if so, how nonseparability affects collective behavior. As I explore the aggregated level, my focus is on the characteristics of the finally selected alternative. I am interested in how an alternative must be "composed" to allow for a joint agreement. Do the groups simply follow the theoretical baseline prediction or is there more to it? If so, what else influences the decision-making and how do votes differ among the decision procedures? The answers to these questions provide a first step towards an understanding of the influence nonseparability exerts on (collective) behavior.

The subsequent sections are structured as follows. I proceed cautiously and first evaluate the reliability of my experimental design. Therefore, I assess my results in comparison to previous research. More specifically, I compare in Sec. 5.1 the most conventional part of my design (the pooled decision-making) to the previous work of S&K (2010) on self-interest and fairness in majority decision-making. Only then do I look at all the data of my voting experiment. For the analysis, I establish a reliable benchmark in two steps. First, I derive in Sec. 5.2 the theoretical baseline solution by extending the core concept to all decision procedures. Second, I discuss in Sec. 5.3 the issue of evaluating a collective outcome in general terms, i.e., how one can measure different criteria such as equality, justice, inclusion, etc. This is related to the elementary question of public policy formulated by Buchanan and Tullock (1962) for the trade-off between efficiency and effectiveness in collective decision-making. I derive statistical measure for both categories and formulate associated hypotheses. Using these benchmarks, I examine the influence of nonseparability on the collective results in Sec. 5.4. Here, I start with an assessment of the explanatory power of the baseline prediction.[243] This allows me to explain some but not all of the deviations observed. The comparison continues on the basis of the empirical computation of the efficiency-effectiveness metrics. Finally, I summarize the findings of this chapter in Sec. 5.5.

[243]This section illustrates in detail how I structure my empirical evaluation of the collective results. All statistical methods, analysis and forms of presentation are explained. Later sections will always refer to these descriptions.

5.1 Reliability of the experimental design

I start by looking into the reliability[244] of the experimental design. Therefore, I compare my findings with previous laboratory work on collective decision-making. As I adapt the general experimental environment presented by S&K (2010), I contrast my results with their findings.[245] The identical aspects between the designs are: a group of individuals decides on distributing points among them; majority rule; subjects interacted over multiple rounds; groups were randomly reshuffled after every round; ballots were repeated until the majority threshold was reached and all previous votes of a round were common knowledge.

However, there are also some aspects which differ significantly. S&K (2010) used five-person committees which voted on eight alternatives while I used six-person committees and offered nine alternatives. These deviations are due to my paramount research goal. The number of six players enabled me to split-up the overall group into two delegations of three players. Nine alternatives enabled me to disperse them into three subsets containing three alternatives each. These modifications of the design affected the probability of reaching a collective agreement. This probability P of agreeing at all in a one-shot majority vote (if players cast their vote randomly) was calculated as

$$P = \left(\frac{1}{n_{altervatives}}\right)^{majority\ treshold} * n_{altervatives} * n_{min\ winning\ coalitions} \tag{5.1}$$

Thus, it depends foremost on the number of alternatives $n_{alternatives}$. The chance that an actor chooses an alternative is just $\frac{1}{n_{alternatives}}$. The *majority treshold* determines the necessary number of actors who have to vote for the same alternative. This can be any of the feasible alternatives and the collective agreement can be reached by any of the possible minimum winning coalitions whose number was given by $n_{min\ winning\ coalitions} = \begin{pmatrix} n_{actors} \\ majority\ treshold \end{pmatrix}$. Thus, an absolute majority vote between five actors has $\begin{pmatrix} 5 \\ 3 \end{pmatrix} = 10$ and a ballot of six actors has $\begin{pmatrix} 6 \\ 4 \end{pmatrix} = 15$ minimum winning coalitions.

For five players voting on eight alternatives this results in a probability of 15.63% for agreeing in a one-shot majority vote. It drops by a factor of 7.59 to 2.06% for six players holding a ballot on nine alternatives. Of course, in both experiments ballots were held repeatedly until an agreement was reached. But the drop in probability documents that the task of reaching a compromise was more difficult in my experimental design.

[244] I follow the definition of Carmines and Zeller (1979, p. 11), according to which reliability is understood as "the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials".

[245] This follows the idea of Guttman (1945) for test-retest reliability which emphasizes that in "dealing with empirical data in any field, the question should be raised: if the experiment were to be repeated, how much variation would there be in the results?" (Guttman, 1945, p. 256)

Another difference between the experiments is that all payoff tables of S&K (2010) contained a core alternative[246]. In contrast, I explicitly designed some of my tables without such equilibrium. This is related to the different research goals of the two experiments. S&K (2010) focused on the performance of the core in majority decision-making. More precisely, they were interested in its robustness against social preferences. I also looked into this question but did not stop with it. In my analysis the measurement of core stability was not the main research interest. Therefore, I used different payoff tables, some with and some without a core. I also altered other characteristic of the tables (cf. Sec. 4.2.4) which eventually resulted in a variety of decision problems.

As already mentioned, S&K (2010) were interested in the relevance of social preferences. They specified a two-way error components model for the selection probability of the core; more precisely, a multilevel mixed-effects logistic regression (S&K, 2010, p. 675). The error component of their model is clustered within each session and within each round of the experiment. In their most simple model the independent variables comprise only the stability of the core alternative. This stability is measured by a monotonic transformation of an Equity, Reciprocity and Competition (ERC) utility function.

The ERC model was introduced by Bolton and Ockenfels (2000). It includes social preferences in an individual's utility function. Specifically, it implies that a player's utility decreases the more the player's own payoff deviates from the mean payoff in their group. The ERC utility function sets a possible trade-off between self-interest and a concern for equity as shown in Eqn. 5.2.

$$U_{ij}(y_{ij}, \sigma_{ij}) = a_i y_{ij} - b_i \left(\sigma_{ij} - \frac{1}{n}\right)^2 \tag{5.2}$$

The utility U of individual i as part of a group of n at alternative j is determined by two components: their absolute payoff y_{ij} and their relative payoff. This relative payoff is given by the difference of their share $\sigma_{ij} = \frac{y_{ij}}{\sum_{j=1}^{n} y_{ij}}$ to the average (i.e., globally equal) share $\frac{1}{n}$. The ratio of $\frac{a_i}{b_i}$ indicates the relative importance player i places on their payment compared to their regard for equality. The theory expects that a_i and $b_i > 0$ which precludes all types of players who prefer inequality, both altruism and envy. Previous applications of the ERC model to experimental data supported these expectations (e.g., Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999).

S&K (2010) incorporated an ordinal stability ranking into the regression analysis in the form of dummy variables, an approach first proposed by Walter et al. (1987). The dummies resembled the log odds ratio between the levels of stability. A coefficient corresponded to the change in the probabilities of the selection of the core if the stability decreased to the next lower level. The ranking indicated for every payoff table for which

[246] S&K (2010, p. 671) stated that all their tables "contain a unique core alternative under the assumption of rationally acting and egoistically motivated committee members."

values of b the ERC utility function predicted a change of the core alternative. The authors included seven dummies dividing the whole value range of b into eight intervals (from $b = 217$ up to $b = 2000$). In their regression they found that, as expected, all significant coefficients of the dummies had negative signs and "infer that the probability of the selection of the core increases monotonically with the core's stability" (S&K, 2010, p. 677).

I replicated the analysis of S&K (2010) to ensure the general reliability of my experimental design. As the first step, I re-estimated their regression model using all 480 observation of their experiment.[247] I modeled the two-way error components as random coefficients for the rounds and for the sessions of the experiment. Tab. 5.1 shows in detail the original results as well as my replication. My calculations mirrored theirs almost entirely.

Table 5.1: Two-way error components model

EXPLANATORY NOTE
The table shows the results of the replication of the S&K (2010) two-way error components model. Statistically significant (two-tailed) at the 0.1 level * and at the 0.01 level ***, SE in parentheses.

N = 480	Dependent Variable: SELECTION OF THE CORE	
	Values reported by S&K (2010, p. 676, table 3)	Own calculations
FIXED PART		
constant	1.89***	1.88***
	(0.38)	(0.37)
Stability217	-1.52***	-1.49***
	(0.43)	(0.42)
Stability403	-1.62***	-1.63***
	(0.41)	(0.41)
Stability625	0.55	0.55
	(0.48)	(0.48)
Stability841	-1.23*	-1.23*
	(0.68)	(0.68)
Stability999	0.55	0.58
	(0.72)	(0.71)
Stability1512	-1.55	-1.58
	(1.14)	(1.13)
Stability2000	1.52	1.51
	(1.11)	(1.11)
RANDOM PART		
Residual SD between sessions	0.25	1.019
Residual SD between rounds	0.73	0.22
Log likelihood	-218.61	-218.86

Next, I applied the two-way error components model for the selection probability of the core to my own results. Due to the different research goals, not all data obtained meets the S&K (2010) criterion for a deterministic equilibrium. Thus, not every procedure or payoff table of my experimental sessions was suitable for a comparison and the number of observations for the regression was significantly smaller. More specifically, I used all collective decisions made on payoff tables with a core alternative under POL; this left me with 58 observations. I accounted for the lower amount of observations by implementing only two dummy variables, which resulted in three intervals of ERC stability: low, medium and high.[248] Payoff tables with low stability deviated from the core for $b \leq 600$,

[247] I am very grateful to Jan Sauermann (University of Cologne) for providing me with the data.
[248] I also considered models with more than three intervals. Yet, when specifying more than two dummy

medium stability included tables which deviate for $600 < b \leq 1200$ and high stability implied an unchanged core for $b > 1200$. This approach was not different from S&K (2010) but less exact. Tab. 5.2 presents the results.

Table 5.2: Selection of the core

EXPLANATORY NOTE
The table shows the estimates of the two-way error components model for the selection of the core. Statistically significant (two-tailed) at the 0.1 level *, SE in parentheses.

N = 58	Dependent Variable: SELECTION OF THE CORE
FIXED PART	
constant	1.099
	(0.816)
Stability600	-0.073
	(0.603)
Stability1200	-1.645*
	(1.11)
RANDOM PART	
Residual SD between sessions	0.342
Residual SD between rounds	0.819
Log likelihood	-37.162

Due to the small number of observations the results should be considered with caution. Nevertheless, they were in line with the findings of S&K (2010). A lower stability led to a lower probability of selecting the core. This correspondence of my most conventional treatment with previous research indicates the reliability of my experimental design.

5.2 Theoretical benchmark predictions

In this section I establish the theoretical baseline for the subsequent empirical analysis of the collective results. Sec. 4.2.4 argues that the appropriate baseline equilibrium is the concept of the core as in S&K (2010). A core is characterized as a unique cooperative solution under the assumption of rationally acting individuals.[249] It follows the deterministic definition of being an alternative for which no single player or subgroup has an incentive to leave the coalition supporting it (Peleg and Sudhoelter, 2003, Chap. 3 and 12). Hence, the core alternative beats all other alternatives in pair-wise comparison.[250]

In relation to previous experiments my design introduced complex decision-making procedures. For example, consider the lower probability of agreeing in a one-shot majority vote in my experiment compared to S&K (2010). As discussed above, the differences

variables, the optimization procedure (STATA's xtmelogit) omitted the additional variable because of collinearity. Thus, the small number of observations restricted the regression to two independent variables.

[249] More specifically, the core concept "satisfies individual, coalition, and collective rationality, inasmuch as it includes only divisions of the payoff such that the players receive at least as much as they could guarantee for themselves by acting independently, every proper subset of the players receives at least as much as it could guarantee for itself by acting together, and the totality of players receives at least as much as it could guarantee for itself by acting collectively as a grand coalition" (Colman, 2003, p. 144).

[250] This section investigates the core performance on the collective level. In Sec. 6.3 I introduce and analyze a probabilistic extension of the core on the individual level.

arose because I chose a group of six instead of five subjects as well as nine instead of eight alternatives. But not only have these larger numbers made it more complex. The splitting of decision competences and the alignment to sequential voting introduced completely new dimensions into the game, e.g., a high level of uncertainty (cf. Tab. 4.2).

RATIONALITY OF GROUP DECISIONS

This higher complexity of the decision task made the baseline analysis even more interesting. Did the game-theoretical concept perform well even under these difficult conditions? A distinct argument in favor is the recurrent observation that team decisions are typically closer to standard game theoretic predictions than individual decisions (Cooper and Kagel, 2005).[251] In their election experiment Forsythe et al. (1996, p. 375) concluded that, overall, "voters cast votes strategically and in equilibrium consistent manners." In accordance, Bornstein and Yaniv (1998, p. 106) found "that groups are more rational than individuals" and Sutter (2008, p. 3) noted that this applies to "a broad variety of strategic- and non-strategic tasks".

An often cited reason for the differences in behavior of groups and individuals is that interaction within the groups affects the actions taken at the intra-group level (Putnam, 1988). Also, the groups can use the knowledge and skills of several people and thus overcome individual cognitive shortcomings (Gigerenzer and Gaissmaier, 2011).[252] This argument is strongly advocated by Surowiecki (2004) who traced the inferiority of individual decisions in various fields back to the aggregation of information within groups.[253] Recent contributions investigated whether behavioral differences can be explained by the fact that motivation differs for individuals and groups. Using laboratory experiments, Kocher and Sutter (2005) found that groups are less trusting than individuals. This is in line with explanations of social psychology[254] which argue that the discontinuity effect between individual and collective action (Schopler and Insko, 1992) is based on the blurred accountability in collectives (cf. Sec. 6.1.4), mutual encouragement of antisocial actions within a group (Wildschut et al., 2001, 2003), the fear of others (Insko et al., 1990) and the reluctance to trust interacting groups (Kugler et al., 2007). All these aspects influence behavior in a way that greed is increased and cooperation diminished. Sauermann (2012, p. 100) summarized this as "compared to individuals the behavior of groups is approaching the model of the Homo economicus".[255]

[251] Cf. Levine and Moreland (1998) for a comprehensive literature overview of the differences of individual and group behavior.

[252] In my experiment subjects could not interact or decide jointly in consultation. Yet, the collective choice procedure required nevertheless an agreement of individual choices. At least the voting process, i.e. the repeated ballots within a round, provided the opportunity for mutual coordination.

[253] Cf. Demertzis (2009) and Mannes (2009) for current contributions to the undecided debate on the "Wisdom of the Crowds".

[254] Cf. Cooper and Kagel (2005) for a comprehensive survey of psychology literature on team versus individual play.

[255] In his study Sauermann (2012, p. 100) supplemented the ERC model for collective actors and found that social preferences might be another explanation for behavioral differences of individuals and groups.

5.2.1 Derivation of the credible core

Sec. 4.4 discusses the data structure of the observations according to whether a procedure resembled the first or second decision rule within a session and if a constant-sum and non-constant-sum table was used. Yet, so far I have omitted the characteristic if a payoff table contained an equilibrium.

The "classical" core concept (Gillies, 1959) is only applicable to POL because standard cooperative game theory studies mostly static one-shot situations (for an overview cf. Peleg and Sudhoelter, 2003). This did not correspond to SIM and SEQ, which exhibited a dynamic and successive decision-making.[256] Yet, the core concept is not limited to a specific institutional rule (for a bicameral extension cf. Bottom et al., 2000).[257] For reasons of consistency I aimed for the same theoretical baseline for all four procedures. Thus, I adjusted the necessary definitions and assumptions of the core to a dynamic game.[258] This did not change my intention to apply cooperative game theory or the fundamental logic behind the core concept. In all procedures the core consisted of the allocation for which no individual or subgroup within the coalition voting for this alternative can do better by deserting the coalition (Peleg and Sudhoelter, 2003).[259]

Research on dynamic or repeated cooperative game theory has been sparse compared to the work done in non-cooperative settings. Only recently has this field developed some promising enhancements for a more general and applicable use of cooperative game theory. A recent overview addressing the core and other cooperative game theory concepts like the bargaining set or coalition formation in repeated, iterated or dynamic settings can be found in Lehrer and Scarsini (2011).

When looking into cooperative game theory by analyzing international pollution control strategies, Dockner and Long (1993) pointed out that a cooperative approach may lead to the first-best solution. But this also requires a high degree of commitment of every participants "that is not likely to be feasible in practice" (Dockner and Long, 1993, p. 16). When such commitment cannot be ensured, an equilibrium solution has to rely on individual strategies to promote cooperation. The crucial question is whether single players (or subgroups) have any incentive to deviate from an arrangement. If yes, this constitutes no stable equilibrium. Only self-enforcing agreements (respectively strategies) can lead to cooperative equilibria.

[256] Already Kramer (1972, p. 171) pointed out that "in general the sophisticated outcome [...] is sensitive to the voting procedure, or order of voting, adopted".

[257] Bottom et al. (2000) derived a bicameral core by using median lines in two-dimensional policy space. Testing their predictions empirically in deliberation experiments the authors recognized a clustering of results around the core areas.

[258] My dynamic game deviates from the setting discussed by Long (2010) as the "state of the system" (my payoff table) did not change over time (within a round). The tables were altered between rounds; i.e., they stayed the same until a collective decision was reached. So, my experiment did not include a "transition equation" (Long, 2010, p. 4). This ruled out open-loop Nash equilibrium (OLNE), Markov-perfect Nash equilibrium (MPNE) or feedback Nash equilibrium (FBNE) applications (cf. Long, 2010, p. 4).

[259] A condensed version of this derivation is part of Fleig and Finke (2013, appendix 3).

A major concern was if in dynamic games without contractual agreements cooperation between players is possible. Here, Tolwinski (2003) showed that negotiated agreements can be enforced by suitably defined strategies. Equilibria have to satisfy the principle of optimality along the (temporal) equilibrium trajectory. Thus, cooperative game theory can constitute the theoretical baseline as long as the solution concept is implemented based on self-enforcing individual strategies. These strategies must account for the dynamic and successive characteristics of the game; in my setting this referred to the division into two delegations as well as the voting sequence.

Next, I had to consider the separated voting of delegations. When studying the core of a finitely repeated discounted cooperative game, Oviedo (2000) re-defined the repeated cooperative game as a repeated game where in each round the agents play a cooperative game. Most importantly, he proved that the core of a repeated cooperative game contains the core of the original cooperative game. I used this idea to account for divided decision-making competencies under delegation. Accordingly, I treated SIM and SEQ as a repeated game, where in each stage (decision of column and row delegation) the participants played a cooperative game.

Kranich et al. (2005) considered the finite horizon of predetermined games and studied three different core concepts: the classical core, the strong sequential core and the weak sequential core. These concepts vary in the degree to which they take the temporal structure of the game into account. In a similar way Becker and Chakrabarti (1995) looked at infinite horizon capital allocation models. They defined the recursive core depending on previous decisions up to that period. These approaches followed earlier work of Bernheim et al. (1987a) and Bernheim and Whinston (1987b) on coalition-proof Nash equilibria. Lehrer and Scarsini (2011) pursued this idea further and identify the credible core as the implementation of a subgame-perfect equilibrium under cooperative game theory. Ray (2007) reached a similar result when incorporating aspects of cooperative and non-cooperative game theory towards a broader definition of the core. He referred to the temporal structure of the game as determined by the farsightedness of a coalition when evaluating the credibility of the core.[260]

To summarize, an equilibrium concept has to be based on self-enforcing individual strategies. This does not affect or prevent my decision to apply cooperative game theory. A repeated cooperative game can be defined as a repeated game, where in each round a cooperative game is played. However, the sequential structure of a game must be taken into account. As with standard theory on strategic games, all "sequential-move games require players to consider the future consequences of their current moves before choosing their actions" (Dixit et al., 2009, p. 79). Here, the equilibrium solution is obtained through backward induction as "rollback equilibrium" (Dixit et al., 2009, p. 79).[261] I used

[260]Credibility represents a central leitmotiv of both cooperative and non-cooperative game theory (Gul, 1997).

[261]This corresponds to the extension of Shepsle's (1979) structure-induced equilibria with perfect-foresight expectations by Denzau and Mackay (1981). Such expectations allow the identification of conditionalities between sequential played issues.

this insight to account for the effects of successive decision-making under delegation. The stability conditions for the core had to apply throughout the entire game against all credible deviations. Therefore, the sequence of coalitions was taken into account to identify a subgame-stable order (Hellman, 2008) which led to the *credible core* as theoretical equilibrium prediction.[262]

SEQUENTIAL DELEGATION PROCEDURE

For SEQ the credible core was identified through a two-stage backwards induction. In each stage the subjects played a cooperative game. I applied the classical core concept to ensure that no individual had an incentive to deviate from the solution found.

First, I identified for the second stage whether a core alternative existed in any of the three sets. If, and only if, all three sets contained a core alternative, this constituted the hypothetical outcome of the collective decision if the first stage agreed on the corresponding set. The second delegation had no incentive to deviate from the core of a set when finally confronted with it. This led to a "reduced" payoff table for the first stage. From the original nine alternatives only three were left (each of the three sets now contained only the left-over core alternative). If, and only if, this "reduced" decision problem of the first stage revealed a core alternative, it represented the solution of the dynamic setting, i.e., the credible core.

For clarification, please imagine a decision under SEQ. First, a delegation decides on the column and then a second delegation on the row. If the first stage delegation chooses set $\{A\}$ the second stage opts for row $\{3\}$, which gives the collective outcome of $\{A3\}$. If the first stage delegation would choose set $\{B\}$, the second stage would vote for row $\{2\}$, and if the first stage delegation agreed on set $\{C\}$, the second stage would settle for row $\{3\}$. This leaves $\{A3, B2, C3\}$ as possible outcomes. These three alternatives constitute the "reduced" decision problem for the first stage. If a comparison of these three alternatives for the first stage members reveals that a core alternative exists, this constitutes the credible core of the complete decision-making procedure.

SIMULTANEOUS DELEGATION PROCEDURE

Under SIM the two delegations did not vote in a predefined sequence. Thus, nor can the stages of the solution concept follow a given order. I implemented this indecisiveness by mirroring the decision-making process. This gave me a second game tree in reversed order. No tree or one or both trees might have provided predictions for a core. But due to the level of uncertainty a stable equilibrium occurred if, and only if, it was the same for both processes, i.e., both sequences predicted the same core. This alternative constituted then the mutual best response of the delegations to each other; i.e., a subgame-perfect equilibrium under cooperative game theory.

[262]Crawford and Haller (1990) showed that even under uncertainty the concept of subgame-perfect equilibrium holds also for repeated coordination games.

For clarification, please imagine a decision in my experiment. Analogous to SEQ, the first stage decides on the column and the second stage on the row of a payoff table. The equilibrium result of this process is alternative $\{B2\}$. If, and only if, a first stage now deciding on the row and a second stage choosing the column leads to the same prediction, then this alternative constitutes the credible core. If the second prediction differs from the first, then this payoff table does not contain a stable equilibrium under SIM.

Since these descriptions of the equilibrium prediction under SEQ and SIM are somewhat abstract, in the following I illustrate the core concept across the three different voting rules with a concrete example. Fig. 5.1 shows a payoff distribution that was actually used in my experiment in all decision-making procedures (cf. Tab. A.7). The payoff table represents a non-constant-sum game and contains, under the assumption of rational and self-interested actors, a unique equilibrium prediction under POL, SEQ and SIM.

Figure 5.1: Example of the determination of the core alternative

EXPLANATORY NOTE
The figure shows payoff table 13 as used in my experiment (cf. Sec. A.8). The equilibrium predictions for this non-constant-sum table are $A1$ under POL and $C3$ under SIM as well as SEQ.

		A	B	C
1	Player 1	22	9	14
	Player 2	16	11	10
	Player 3	30	6	14
	Player 4	17	2	42
	Player 5	3	21	9
	Player 6	22	17	18
2	Player 1	5	20	33
	Player 2	12	41	1
	Player 3	33	7	8
	Player 4	15	14	1
	Player 5	14	22	38
	Player 6	19	18	10
3	Player 1	8	7	20
	Player 2	4	26	14
	Player 3	19	17	13
	Player 4	10	12	30
	Player 5	42	19	1
	Player 6	20	16	32

Under POL $\{A1\}$ constitutes the core. This alternative beats all other eight alternatives in a pairwise comparison by majority vote.[263]

Under SEQ the group of six players is divided into two delegations with players 4, 5 and 6 voting first on the column and players 1, 2 and 3 voting afterwards on the row of the collective outcome. The credible core is identified through backwards induction. If the first stage delegation would choose set $\{A\}$, the second stage would opt for row $\{1\}$ (by the votes of players 1 and 2). If the first stage delegation would choose set $\{B\}$, the second stage would vote for row $\{2\}$ (by the votes of players 1 and 2). If the first stage delegation would choose set $\{C\}$, the second stage would settle for row $\{3\}$ (by the

[263] Alternative $\{A1\}$ wins the pairwise comparisons according to the following winning coalitions: against $\{A2\}$ by the votes of players 1, 2, 4 and 6; against $\{A3\}$ by the votes of players 1, 2, 3, 4 and 6; against $\{B1\}$ by the votes of players 1, 2, 3, 4 and 6; against $\{B2\}$ by the votes of players 1, 3, 4 and 6; against $\{B3\}$ by the votes of players 1, 3, 4 and 6; against $\{C1\}$ by the votes of players 1, 2, 3 and 6; against $\{C2\}$ by the votes of players 2, 3, 4 and 6; against $\{C3\}$ by the votes of players 1, 2, 3 and 5.

votes of players 1 and 2). This leaves $\{A1, B2, C3\}$ as possible outcomes which the first stage delegation has to consider ("reduced" decision problem). With two votes in favor (players 4 and 6) and one vote against (player 5) then $\{C3\}$ beats both $\{A1\}$ and $\{B2\}$. Thus, $\{C3\}$ constitutes the credible core under SEQ.

Under SIM the reversed order of the two stages is also considered. Thus, the group of six players is again divided into two delegations but players 1, 2 and 3 vote first on the row and players 4, 5 and 6 vote afterwards on the column of the collective outcome. If the first stage delegation would choose set $\{1\}$, the second stage would opt for column $\{C\}$ (by the votes of players 4 and 6). If the first stage delegation would choose set $\{2\}$, the second stage would vote for column $\{A\}$ (by the votes of players 4 and 6). If the first stage delegation would choose set $\{3\}$, the second stage would settle for column $\{C\}$ (by the votes of players 4 and 6). This leaves $\{C1, A2, C3\}$ as possible outcomes which the first stage delegation has to consider (again a "reduced" decision problem). Next, with two to one votes $\{C3\}$ beats both $\{A2\}$ and $\{C1\}$ (players 1 and 2 in favor). Thus, as the equilibrium predictions for both stage orders match, $\{C3\}$ constitutes the credible core under SIM.

5.2.2 Distribution of the equilibrium solution over the payoff tables

The adaptation of the core concept for the two delegation procedures led to the distribution of payoff table characteristics described in Sec. A.9. Tab. A.6 shows that for each procedure at least four payoff tables existed which corresponded to every possible table property combination - with and without a core alternative as well as constant-sum and non-constant-sum game. Due to the differences in determining the (credible) core the table assignment to the categories differed between procedures. Above all, this concerned whether a table contained a core at all. And if so, it also affected the specific equilibrium within a payoff table between procedures.[264]

More important than the general structure alone is the actual frequency of use of each table. Here, as shown in Sec. A.10, the random assignment rule resulted in, as expected, an uneven usage of payoff tables. The interaction of these two elements resulted in the data structure of all collected observations of my experiment. Sec. A.11 discusses the number of observations with respect to the distribution of the table characteristics. For now, I limit the analysis to the equilibrium prediction. Thus, Tab. 5.3 displays the frequency of observations according to the existence of a deterministic core alternative. During each procedure I investigated payoff tables with and without equilibrium to observe decision-making under a variety of configurations (cf. Sec. 4.2.4).

[264]This is in line with Hammond and Miller (1987) who demonstrated that bicameralism can create core solutions in settings where simple majority rule is affected by cycling among outcomes and vice versa. In Sec. A.8 I indicate below every payoff table if it constituted a constant-sum or non-constant-sum game and the respective core alternative for every procedure (if one exists).

Table 5.3: Data structure according to the existence of an equilibrium alternative

EXPLANATORY NOTE
The table depicts the number of observations in terms of collective decisions made in the experimental sessions. The observations are divided according to procedure and whether or not the payoff table used contained a core alternative.

		CORE ALTERNATIVE		Σ
		does exist	does not exist	
PROCEDURE	pooling	58	18	76
	simultaneous delegation	42	43	85
	sequential delegation	48	12	60
	Σ	148	73	221

Overall, I conducted more experimental rounds with tables containing a core solution (148 to 73). Looking into the procedures, the distribution of payoff tables with and without a core showed equal amounts under SIM (42 to 43). For POL and SEQ many more observations using tables with a core alternative were collected (under both rules in around 75% of the cases). Summing up, for the performance analysis of my theoretical baseline this gave me 58 collective observations under POL as well as 42 under SIM and 48 under SEQ.

5.3 How to judge a collective decision

My experiment contrasted the situation in which the collective decision is reached in the group as a whole to the situation in which the group is split into two delegations which must decide simultaneously or sequentially. The most intriguing aspect of this approach was the opportunity to compare the behavior and outcomes of the joint and the separate game structure within one experimental design.

It also distinguished my experiment from previous research which investigated the implications of different decision rules. I discuss this literature in Sec. 5.3.1 and define statistical measures for evaluating the chosen alternative out of and according to it. Next, Sec. 5.3.2 derives corresponding hypotheses for the collective results. The specific operationalization of the metrics used is explained in Sec. 5.3.3. These statistical measures served, in addition to the theoretical equilibrium, as the second starting point for the assessment of the selected alternatives. Only by establishing this credible classification was the analysis able to reach conclusive findings.

5.3.1 Decisions costs and welfare effects

Laws and regulations coordinate the various activities of individuals within every society. Such rules are of paramount importance for all individual and collective decision-making.[265] Brennan and Buchanan (2000) defined the investigation and evaluation of social rules as the subject matter of modern political economy. Since Buchanan and Tullock (1962) various contributions have contrasted the pros and cons of different k-majority

[265] For a more comprehensive overview cf. Mueller (2003, part II, p. 67-208).

rules[266]; most prominent are unanimity and absolute majority rule[267]. This also includes studies which introduced veto players as a key element instead of different voting thresholds (e.g., Chen and Ordeshook, 1998; Kagel et al., 2010) as the implications are the same as under unanimity.[268]

Current research does not provide a clear-cut answer for the search of the optimal social choice mechanism.[269] For example, investigating jury decision rules Guarnaschelli et al. (2000, p. 407) found "fewer outcomes analogous to incorrect convictions under unanimity rule than under majority rule"; certainly a positive effect. Consistently, many theoretical computations led to the conclusion that "for outcomes that are both Pareto optimal and Pareto superior, unanimity rule outperforms majority rule" (Dougherty and Edward, 2012, p. 655). Colomer (2001) discussed different voting rules such as unanimity, majority and plurality vote as well as various types of political regimes. He investigated parliamentarism with either majority or proportional representation and also presidentialism for its social efficiency. In the end he concluded that unanimity rule equilibria are Pareto optimal and, maybe even more striking, that majority rule equilibria rarely exist. The Pareto argument is further illustrated by Mueller (2003). In its very elementary logic unanimity ensures that only Pareto improvements are possible because otherwise a member would contradict their own interests.

Yet, these theoretical findings are weakened as they not hold up in simulations which use "random proposals and sincere voting" (Dougherty and Edward, 2012, p. 662). Also, Colomer (1999, p. 543) found that unanimity decisions heavily "depend on the initial state or the status quo. The closer the status quo is to the ideal points of the actors, the more restricted, more biased and likely less socially efficient the set of decisions by unanimity tends to be." Such a categorization implies that general statements of superiority should be avoided.

Contrary to the theoretic assessment that "unanimity rule is the only voting rule certain to lead to Pareto optimality" (Johnson, 1991, p. 161), multiple experimental evidence suggests "that majority [rule] can produce larger welfare effects" (Sauermann and Glasmann, 2011, p. 373). Colomer (2001, p. 71-73) even came to the exact opposite conclusion, namely that majority rule is, indeed, better at obtaining Pareto-optimal outcomes than unanimity rule. Sauermann and Glasmann (2011, p. 391) attributed this to the fact that when

[266] Dougherty et al. (2009, p. 1) defined a k-majority rule to require at least k from a total of N individuals to vote in favor while it holds that $\frac{N}{2} < k \leq N$. Extensions to the cases of non-voters or "votes to abstain" are discussed in Dougherty and Edward (2004).

[267] Those constitute two special cases of k-majority rule: majority rule for $k = \frac{N}{2}$ and unanimity rule for $k = N$ (Dougherty et al., 2009, p. 1).

[268] The well-known contribution of Tsebelis (2002) clearly demonstrated to what extent and how veto players shape policy.

[269] In addition to unanimity and absolute majority rule this statement also holds true for alternative voting mechanisms. For example, Forsythe et al. (1996) conducted laboratory election experiments under plurality rule, approval voting, and Borda rule. The authors also studied the effects of (non-binding) pre-election polls. Crucial for the search for the optimal electoral system (but unfortunately also disappointing) is "that Condorcet losers occasionally win regardless of the voting rule or presence of polls" (Forsythe et al., 1996, p. 355).

evaluating different democratic voting norms "the majority rule is a strong incentive to cooperate"[270].

In the end, the assessment for the optimal social choice mechanism depends on the decision costs which are set against the welfare effects of the voting rule (Buchanan and Tullock, 1962; Kaiser, 2007a).[271] Miller and Vanberg (2014) directly investigated the effects of different decision rules on the costs of decision-making in a laboratory experiment. They observed more rejections, bullish behavior and costly delays under unanimity rule. This implies "support of less-than-unanimity decision rules" (Miller and Vanberg, 2014, p. 20) in multilateral bargaining situations. Interestingly, Miller and Vanberg (2014) stated that an issue worth exploring in future research is the effect of group size. This way one could test the classic argument of Buchanan and Tullock (1962), which states that the costs of decision-making increase with the size of the decision-making body.[272] This is feasible in my design when looking at groups of six as well as three members.

I followed the contributions discussed by evaluating the collective majority decisions based on decision costs and welfare effects. This represents a trade-off between (cost) efficiency and (welfare) effectiveness. Those two categories were evaluated using five key figures which I derived from the literature on collective decisions: *decision-making efficiency*, *social welfare allocation*, *distribution of wealth*, *approval rate* and *stability*. Subsequently, I explain the origin and intention of each metric.

EFFICIENCY

First, I look at the efficiency aspect. As many collective decisions are typically time-consuming, its distinction is straight forward. Efficiency depends on the ability to reach a decision in a parsimonious way, i.e., to minimize transaction costs (Coase, 1937, 1960). For clarification, in my experiment neither the sum nor the distribution of points of an alternative is altered within a round across the single ballots. This distinguishes my conception of efficiency from experiments where the available amount of points decreases with the number of ballots.[273] Instead, it corresponds to the literature on management theories in which "efficiency is the achievement of the ends with the least amount of resources" (Olum, 2004, p. 6). In other words, to solve the collective agreement problem in as few steps as possible. I designate this measure *decision-making efficiency*.

[270] The work of Dougherty et al. (2009) showed yet another possible explanation, the specific experimental setting. The authors concluded "that unanimity rule may not be particularly adept at selecting Pareto optimal outcomes if the starting point is not in equilibrium" (Dougherty et al., 2009, p. 23).

[271] This consideration goes far back to Wicksell (1896) and his assessment of the Swedish tax system.

[272] A non-agreement on a topic is generally seen as disadvantage or waste of time. But under some circumstances it may, however, be desirable. For example, in their jury experiments Guarnaschelli et al. (2000, p. 419) found that "larger juries may convict fewer innocent defendants than smaller juries under unanimity."

[273] For example, in the bargaining experiment of Miller and Vanberg (2014) a group takes a vote on a proposed payoff distribution of an initial endowment. "If a simple majority accepts the proposal, the game ends and each player receives his allocated amount. If not, the pie shrinks by a certain factor and a new round begins. Thus, the costs of bargaining consist of the lost surplus" (Miller and Vanberg, 2014, p. 6).

EFFECTIVENESS

Next, I turn toward statistical measures for their effectiveness. The range of possible criteria for assessing welfare effects is larger and more diversified than for efficiency. Most clearly relevant and immediately apparent is the amount of allocated points which results in the *social welfare allocation* of the subjects.

In addition, not only the sheer scale of welfare[274] but also its distribution is relevant (Rawls, 1971). Already Black (1948, p. 29) emphasized the "persistence of disharmony and discord" in collective decisions. More often than not the equilibrium solution predicts a highly unequal distribution of outcomes in which the majority cares little about the minority's well-being (Saunders, 2010b). Thus, I also consider the *distribution of wealth*. Interestingly, previous experiments have proved that distribution issues are related to the classification of participants.[275] Allocating subjects to be a "row" or "column" player, as in my design, influences preferences (Charness et al., 2007b; Goette et al., 2006). For example, Chen and Li (2009) measured higher earnings for ingroup matching in their allocation games. Banuri et al. (2011) investigated the reasons for nepotism[276] and the influence of anti-nepotism laws. They found that prohibitions reduce trust and that, with salient membership, subjects engaged in socially costly nepotism. More generally, whether the resulting effect was socially desirable or not (i.e., whether it led to more social welfare) depended on the particular environment or type of the game. Yet, decisive was always the salience of membership, i.e., the subjects' identification with their assigned group (Charness et al., 2007b, p. 1362).[277]

As well as the realized outcome itself, its emergence and persistence also have to be considered. In a general context, participation and inclusion are desired and necessary for a functioning democracy (Schäfers and Zimmermann, 2005).[278] This question refers to research on coalition-formation (e.g., Schofield, 1996) and is interesting with respect to differences in the size of winning coalitions across procedures. S&K (2010), Miller and Vanberg (2014) as well as Sauermann and Glasmann (2011) found that people make use of their veto rights. It seemed reasonable to expect that subjects use their opportunities for excluding supernumerary players when aiming for the majority threshold. Thus, groups

[274] In terms of my experiment, scale refers to the amount of points accumulated by the group when reaching a collective agreement.

[275] Research on (group) identity is a wide field. Regarding the methodological framework, Sutter (2008) extended the findings on group membership of Charness et al. (2007b) who had focused on strategic games also to non-strategic settings. Identity is also a fairly universal pattern. In general, Fehr et al. (2008) showed that the understanding of identity already develops between the ages of 3 to 8 years. More specific studies investigated how personal relations affect the relation between managers and employees (Brandts and Sola, 2006) or looked at identity when making investment decision (Güth et al., 2005).

[276] Banuri et al. (2011, p. 1) defined nepotism as "discrimination in favor" of a group member relative to the population (cf. Becker, 1971).

[277] Brewer (1999) advocated a distinction between ingroup favoritism and outgroup discrimination. Yet, my design does not allow distinguishing between them. Thus, I was not able to determine whether there is affection towards ingroup or rejection of outgroup subjects.

[278] Admittedly, in my laboratory experiment such a value judgment was less clear. That all participants go along is no common wisdom for such an abstract context.

should opt for small coalitions to enforce beneficial alternatives (for a literature overview on games with equilibria in minimal winning coalitions cf. Chen and Ordeshook, 1998). This is supported by experimental findings. For example, Berl et al. (1976, p. 473) observed minimal winning coalitions "four out of five times" in their collective voting and van de Kragt et al. (1983) recorded predominantly "minimal contributing sets" in their public good experiment. Taking this into account, I examined the *approval rate* of every collective agreement. With respect to persistence, I considered the *stability* among the different decision procedures. It is an important purpose of organizational procedures to ensure reliable results (e.g., Shepsle, 1979) and secure durable policies (Colomer, 2001, p. 208). That it is reasonable to avoid cycling in majority decisions is one of the few undisputed understandings in public choice (Shepsle and Cox, 2007).

5.3.2 Hypotheses about the treatment effect

Sec. 2.5 states common arguments on the influence of NSP on decision-making. Furthermore, the section points out that people are subject to cognitive limitations (Oaksford and Chater, 1992; Stanovich and West, 2000) and are not able to consider thoroughly all implications and restrictions of a decision rule. I look more thoroughly into this matter in Chap. 6 when assessing individual voting behavior and variations in cognitive capabilities. For now, it suffices to distinguish between the four decision situations unraveled in Tab. 4.2: POL SIM as well as first and second stage SEQ.

The arguments were formulated in general terms as they focus on basic patterns. In the following I refine them with testable hypotheses according to the concrete design of my experiment. POL, which constitutes my most conventional treatment, serves as a genuine link to prior contributions. Here, a two-dimensional decision problem was submitted to a group deciding collectively. The two dimensions of the decision problem (i.e., row and column of the payoff table) were clearly interlinked. Thus, the basic principle of nonseparability implies that those parts should not be separated. If they are split nonetheless the concept of NSP predicts a "first-mover" advantage for sequential (ARGUMENT 2) and a sub-optimal collective outcome for simultaneous decision-making (ARGUMENT 3). In terms of the decision situations in my experiment this relates to a contrasting juxtaposition of i) first and second stage under SEQ to determine a possible first-mover advantage and of ii) POL and SIM to investigate potentially sub-optimal outcomes.[279] I accomplish and structure this comparison based on the just in Sec. 5.3.1 derived statistical measures. The first two hypotheses[280] aim at social welfare allocation and outcome stability.

[279] ARGUMENT 3 distinguishes between decisions made jointly and separately. This corresponds to the comparison of POL and SIM. Yet, SIM introduced the highest level of both cognitive complexity and uncertainty (cf. Tab. 4.2). Thus, it was more difficult for the participants to fully comprehend the decision problem under SIM than under SEQ. When evaluating the results I therefore also discuss the differences between the two delegation procedures.

[280] The hypotheses follow those two defined in Sec. 3.3.3. Therefore, I continue the number sequence.

HYPOTHESIS 3 (first-mover advantage): Under the sequential delegation procedure the first stage members achieve a higher social welfare and more stability than the second stage participants.

HYPOTHESIS 4 (sub-optimality): The simultaneous delegation procedure achieves a lower social welfare allocation and less stability than the pooling procedure.

Next, I focus on the distribution of wealth and the approval rate of the collective decision. Here, I also have to consider the insights about group identity; under both delegation procedures the subjects were assigned to vote either on the column or the row of the payoff table. Taking first-mover advantage and group identity together, under SEQ the first stage participants should decide in accordance with their mutual benefit. But then the second stage was condemned to fight for the leftovers (especially in the case of a constant-sum table). Thus, consensus should be less common in the second stage because of scarce resources. Also, if the assignment proves salient enough, the participants should have been more willing to compromise under SIM than under POL.[281]

HYPOTHESIS 5 (scarcity): Under the sequential delegation procedure the first stage members achieve a more equal distribution and a higher approval rate than the second stage participants.

HYPOTHESIS 6 (compromise): The simultaneous delegation procedure achieves a more equal distribution and a higher approval rate than the pooling procedure.

In addition to the configuration of a collective decision the classic argument of Buchanan and Tullock (1962) focuses on the swiftness with which it is achieved. They stated that decision costs increase with the size of the electorate. I investigated this group size effect by contrasting decisions between groups of six and groups of three members. Of course, the necessary negotiations for reaching a collective decision depend also on the intricacy of the task. So a simpler problem should lead to a faster solution.

HYPOTHESIS 7 (intricacy): Under the sequential delegation procedure the first stage members achieve a lower decision-making efficiency than the second stage participants.

HYPOTHESIS 8 (swiftness): The simultaneous delegation procedure achieves a higher decision-making efficiency than the pooling procedure.

To summarize, all hypotheses for the evaluation of the collective decisions are listed in Tab. 5.4. This illustrates that the hypotheses 3, 5 and 7 address differences in the performance of first and second stage participants under SEQ. Under SIM both delegations had to perform the exact same task and there was no distinction according to succession. Thus, no differences in performance between these delegations were expected. More

[281] This hypothesis is supported by existing literature on social preferences. Previous studies found that they are stronger among subjects who interact in small groups compared to those in large groups (Fehr and Schmidt, 1999). The question remains whether the reduction from six to three persons is sufficient.

specifically, under SIM the two delegations achieve the same decision-making efficiency, social welfare allocation, inequality distribution, approval rate and stability. I exploited this fact to conduct a robustness check against artificial findings due to, e.g., an unaccounted imbalance in the payoff tables between rows and columns.

Table 5.4: Hypotheses for the treatment effect

EXPLANATORY NOTE
The table summarizes the hypotheses about the collective results. They are structured according to the respective statistical measures. For every hypothesis its designation, the involved decision situations as observation unit and the expectation of their relative performance are given.

HYPOTHESIS		EXPECTATION
		SOCIAL WELFARE ALLOCATION and STABILITY
3	first-mover advantage	under sequential delegation the first stage performs better than the second stage
4	sub-optimality	simultaneous delegation performs worse than pooling
		DISTRIBUTION OF WEALTH and APPROVAL RATE
5	scarcity	under sequential delegation the first stage performs better than the second stage
6	compromise	simultaneous delegation performs better than pooling
		DECISION-MAKING EFFICIENCY
7	intricacy	under sequential delegation the first stage performs worse than the second stage
8	swiftness	simultaneous delegation performs better than pooling

5.3.3 Operationalization of the statistical measures

Before I turn to the empirical analysis in the next section, I describe subsequently the empirical properties and concrete implementation of the five statistical measures.

DECISION-MAKING EFFICIENCY

In my experiment efficiency refers to a delay of the collective decision. This was measured by the number of ballots necessary to finally come to a collective agreement.

SOCIAL WELFARE ALLOCATION

Welfare effects were related to the collectively gathered sum of points. Due to the random assignment rule of the payoff tables (cf. Sec. 4.2.4) I normalized the results. This enabled me to compare different tables and procedures. Eqn. 5.3 shows the corresponding calculation. Every (non-constant-sum) payoff table offers the subjects a continuum of alternatives of which one inhibits the maximal x_{max} and one the minimal x_{min} attainable amount of points for the group. The transformation function sets the actually achieved amount of points x_i in relation to this continuum. It transforms the continuum in a way that the new value range extends for every payoff table from 0 to 100.[282] The transformed amount of points $T(x_i)$ then indicates how successful a decision was in terms of this range. Thus,

[282] The normalization corresponds to a linear scale transformation instead of a z-standardization (Studenmund, 2006, p. 541-544). Through normalization the results from different payoff tables became comparable by conversion into a homogeneous scale. Standardization would, in addition, always lead to mean of zero and a SD of one for the standardized variable (for proof cf. Gujarati and Porter, 2008, p. 173ff and appendix 6A, Sec. 6A.2).

in the analysis each decision's relative allocation performance is reported instead of just presenting the absolute sum of points.[283]

$$T\left(x_i\right) = \frac{100}{x_{max} - x_{min}} \times \left(x_i - x_{min}\right) \tag{5.3}$$

DISTRIBUTION OF WEALTH

I measured distributional aspects using the standard deviation as well as the Gini coefficient (Gini, 1912; Hirschman, 1964) of an alternative. Both are well-known and reliable indicators of inequality (e.g., Bellu and Liberati, 2006; Haughton and Khandker, 2009).

As with the allocation of social welfare its distribution can be expressed in absolute or relative terms. In the following, each decision's relative inequality performance is reported instead of just presenting the absolute amount of standard deviation. For the standard deviation this was calculated in accordance with Eqn. 5.3.[284] The value range of the Gini coefficient is $\left[0; \frac{n-1}{n}\right]$ (Wagschal, 2009, p. 130).[285] In order to roughly keep this range I normalized the coefficient to the interval $[0; 1]$. This was achieved by using 1 instead of 100 in Eqn. 5.3. Here, 0 means a maximal uniform distribution and 1 represents a maximal inequality.[286] Both metrics show how close the committee has come to reaching an equal split. The greater the concentration is, i.e., the imbalance in the selected alternative, the higher the numerical values.

APPROVAL RATE

The extent of agreement for every collective decision was determined endogenously in my design. Groups could agree unanimously on an alternative or enforce a decision on the basis of majority rule. In either way the selected outcome determined the payoffs for all group members. There was no reward or deduction for consensus. I counted the approval rate by the size of the winning coalition (i.e., the number of matching votes) which reached the collective agreement.

[283] Through normalization the results are expressed as a relative success rate. Imagine a collective decision on a payoff table of which the highest sum of points of an alternative equals 80 and of which the lowest amount is 30. A group that manages to select the alternative which adds up to 80 points was completely successful and ends up, according to Eqn. 5.3, with a success rate of $\frac{100}{80-30} \times (80-30) = \frac{100}{50} \times 50 = 100$. On the other hand, if the group ends up with an alternative that adds up to 55 points (i.e., midway between 80 and 30), its success rate is only $\frac{100}{80-30} \times (55-30) = \frac{100}{50} \times 25 = 50$.

[284] Through normalization the results are expressed as relative inequality. Imagine a collective decision on a payoff table of which the largest imbalance is 80 and of which the smallest is 30. A group that manages to select the alternative with an inequality of 30 effectively agrees on the most equal split which, according to Eqn. 5.3, results in a relative inequality of $\frac{100}{80-30} \times (30-30) = \frac{100}{50} \times 0 = 0$. On the other hand, if the group ends up at an alternative with an inequality of 55 points (i.e., midway between 80 and 30), its relative inequality performance is $\frac{100}{80-30} \times (55-30) = \frac{100}{50} \times 25 = 50$.

[285] In the case of my experiment, a group of six players ($n = 6$) decided on the distribution of points. This resulted in a possible range of $\left[0; 0.8\overline{3}\right]$.

[286] The term "maximal" should be interpreted as the maximal possible distribution in this decision in the experiment. For example, if a choice between alternatives provides the Gini coefficients of 0.2, 0.4 and 0.7 then a normalized Gini coefficient of 0 would refer to 0.2 and of 1 to 0.7 for the result.

STABILITY

The robustness of the collective decision was not directly tested in my experiment. Once an agreement was reached, the respective alternative was finalized and not again put to vote (cf. Sec. 4.2.4). I therefore approximated the persistence of outcomes by means of the homogeneity of the collective results. This was measured by how homogeneously the voting pattern within and between treatments was when facing the same decision problem (i.e., the same payoff table).

Summing up, the five statistical measures for the evaluation of the collective decisions were operationalized as follows:

- *Decision-making efficiency* is counted as the number of ballots conducted.
- *Social welfare allocation* is expressed as the relative allocation performance.
- *Distribution of wealth* is measured as the relative inequality performance.
- *Approval rate* is represented by the size of the winning coalition.
- *Stability* is approximated by the homogeneity of the collective results.

TREND ANALYSIS

When one investigates the influence of NSP based on these metrics a static assessment is not sufficient. An examination of the results over time is also necessary (cf. Tab. 4.3). With every round of the experiment the experience of the subjects with the experimental setting grew. While a session progressed, they were trained in the process and knew how reaching an agreement worked. This should be visible, e.g., in terms of fewer ballots per collective decision, a higher social welfare allocation and greater stability in later rounds compared to those at the beginning. Yet, it is not clear after how many rounds a change should occur. Because of the vague context I operationalized this computation by means of a trend analysis. In this way I avoid having to specify a concrete cut-off point. The same consideration applies to first and second decision rule of a session. Thus, I expected a better performance in terms of decision-making efficiency, social welfare allocation, distribution of wealth, approval rate and stability after participants had gained experience.

5.4 Results

This section investigates the collective results of my voting experiment. As first step I look into the performance of my theoretical baseline. In addition to its contribution about collective behavior, this subsection serves an important additional purpose. The discussion of the experimental design has shown that in the empirical assessment some aspects are of particular importance (cf. Sec. 4.3.5). The evaluation had to consider trend analysis, statistical independence of observations, etc.[287] When assessing the theoretical benchmark predictions I discuss the necessary steps of the empirical analysis in detail. I outline the used test statistics and explain the structure of the summary tables which present my results (e.g., observations separated according to be first or second decision rule of a session). Subsequent analyses follow this pattern; I use the same methods, but I do not discuss every aspect in such detail.

In the tradition of Buchanan and Tullock (1962, cf. also Kaiser, 2007) collective decisions resemble a trade-off between decision costs and welfare effects. In other words, one has to choose between efficiency and effectiveness of the reached agreement. Efficiency turns toward the ability to reach a decision in a short and parsimonious way. Here, the consequences of the chosen alternative do not matter. Effectiveness, on the other hand, depends solely on the specifics of the selected option. It is not important how long it took until the decision was made; only its output properties count. This is the second step of my empirical analysis. I look successively at the derived statistical measures and formulated hypotheses (cf. Sec. 5.3.2); this provides a thorough overview on the achievements of the different decision procedures. I also analyzed possible trends of the metrics over time, i.e., over the progressing experimental rounds. This implied a lot of additional tables. To keep the analysis comprehensible I include them in Sec. A.12. This serves solely for the purpose of a concise presentation. All results are discussed and clearly linked to the corresponding table in the appendix.[288]

5.4.1 Selection of the core

In their simpler design S&K (2010, p. 674) found that 69% of all their committee decisions resulted in the selection of the core alternative. As I modify their general experimental environment I considered this value as a reference point. With respect to my data, not all observations were suitable for a comparison. But unlike in Sec. 5.1, I did not have to confine myself solely to POL. I used all collective decisions made on payoff tables with a core alternative. Across all procedures these were 148 observations.

Before comparing the core performance one last modification was necessary. Under SIM the final outcome is determined as the intersection of the two delegations' decisions. Both delegations chose only a subset and not a specific alternative. This is an important difference to SEQ where at least the second stage delegation picked one concrete alternative. Thus, the performance of the core in predicting the collective voting behavior was calculated at the level of the two delegations; i.e., the percentage with which a delegation chose the subset that belonged to the credible core. As every group decision consisted of two delegation decisions the number of observations at this level was twice as large.

[287] Tab. 4.3 summarizes the key characteristics and their implications for the empirical analysis.

[288] Without getting ahead of myself, I have moved the trend analysis into the appendix because it contains only few significant results. The knowledge gained is therefore rather low.

The core performance under each treatment is shown in Tab. 5.5[289]. As discussed before, observations are not statistical independent across rounds (cf. Sec. 4.3.2). Therefore, the unit of observation must be session averages per procedure of the variable in question (cf. Frechette, 2012).

Table 5.5: Core performance

EXPLANATORY NOTE
The table shows the performance of the core in predicting the collective decisions. Each cell indicates the number of correct predictions in relative terms. The observations are separated by procedure and for being the first or second decision rule of a session. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE				first decision rule			second decision rule		
	N	mean	SD	N	mean	SD	N	mean	SD
pooling	9	42.6%	16.4	2	45.8%	5.9	7	41.7%	18.6
simultaneous delegation delegation level	8	40.6%	13.9	6	37.8%	10.4	2	49.0%	25.0
sequential delegation	8	39.6%	27.6	5	30.8%	29.1	3	54.2%	7.2

It is quite apparent that the performance of the core in forecasting the collective decisions was lower in my experiment than in the case of S&K (2010, p. 674). For every procedure, the concept could predict about 40% of the decisions. For SIM this refers to the delegation instead of the group level (as discussed in Sec. 4.3.4), where a core set selection was independent of the other delegation. So, a non-selection of one delegation did not imply a non-selection for both. A performance of about 40% might not be considered as a very good predictor, but it is also not necessarily a poor one. The frequency of selection is clearly above a 11% probability of randomly drawing one of the nine alternatives.

When examining the results across procedures, decision rules, etc., I make my assessments not just by looking on the obtained average values and applying a rule of thumb. Every single statement is based on a nonparametric test (Siegel, 1988). This is either the Wilcoxon matched-pairs signed-ranks test (Wilcoxon, 1945)[290] for within-subject comparisons or the Mann-Whitney U-test (Mann and Whitney, 1947) for between-subject analysis (cf. Sec. 4.2.7). Both are nonparametric rank sum tests and evaluate whether the central tendency of two samples correspond,[291] they differ according to their assumption whether the samples used are related. I stratified my observations according to procedure, group and delegation level, first and second decision rule, as well as under SEQ due to first and second stage. For reasons of clarity I do not report results for every pairwise

[289] For reasons of clarity the presentation of the joint separation for both table properties has been omitted as it contained no additional information.

[290] I used the signed-ranks test (Wilcoxon, 1945) and not the sign test of matched pairs (Arbuthnott, 1710; Snedecor and Cochran, 1989), because the signed-ranks test takes into account not only the positive or negative sign of differences, but also the extent of the differences between the paired samples.

[291] Contrary to a parametric two sample t-test the assumption of a normal distributed dependent variable is not necessary (Gibbons, 1976). This is required as provident tests for normal distribution (Royston, 1992; Shapiro and Wilk, 1965) on almost all obtained variables turn out negative. The violation of this assumption would result in unreliable inferences and misleading interpretations of the t-test (for an assessment when either test cannot be applied due to extremely skewed data cf. McElduff et al., 2010). Although it is general assumed that the t-test is more powerful (because nonparametric tests convert the observed values into ranks) this does not hold for large samples (cf. Motulsky, 2010, Chap. 37).

comparison. Instead, I list all significant findings, including the respective (two-tailed) significance level. All unreported values signal an insignificant test.

Looking more into the details of Tab. 5.5, neither comparison across procedures resulted in a clear pattern. Although the performance values seemed higher for the second decision rule at first glance, the differentiation for decision-rule resulted in no statistically significant difference under POL, SIM or SEQ. This was also true when considering all observations together.

In my experiment subjects participated in multiple rounds in every session. So, in addition to the overall performance of the core also its progression over the rounds is of interest. After the subjects got acquainted to the game (i.e., after some training), it is reasonable to assume that their voting patterns might change (cf. Sec. 4.2.5). If this is true, the performance of the core should be better in later rounds.

I analyzed the development over the single rounds by using STATA's *somersd* package (implemented for STATA by Newson, 2002),[292] a nonparametric rank-statistic.[293] This package performs a maximum likelihood fit to obtain association measures (and CI) of a predictor and the dependent variable. In other words, it depicts the difference between the probability that two variables are concordant and the probability that they are not (Newson, 2013, p. 1). The estimated value D might be interpreted as a "predictor performance indicator" (Newson, 2006a, p. 312). In terms of my analysis, it showed the ability of the experimental round to predict the performance of the core, i.e., if the core alternative was chosen more often in later rounds.

Tab. 5.6 shows the results of the trend analysis on the performance of the core.[294] The observations are split according to procedure and being the first or second decision rule within a session. For none other than two estimates a significant trend was discovered. In both cases the trend occurred within the second decision rule of an experimental session. Thus, the trends were observed only after a procedural change. Merely repeating the game did not alter the core performance in any of the procedures. Only in combination with altering the decision rules did repetition lead the subjects to achieve the equilibrium result.

Under SEQ I found a positive and under SIM I detected a negative trend over the rounds. The finding for SEQ was more intuitive as it also corresponded to the findings in Tab. 5.5. All SEQ rounds as second decision rule were run after playing SIM before (cf. Sec. 4.4). Thus, the procedural change (the introduction of sequential voting) in combination with

[292] The main advantage of this STATA package is that it calculates CI using jacknife variances (Newson, 2006b). More precisely, it computes the CI of rank order statistics calculated by the Wilcoxon rank-sum test (Wilcoxon, 1945). This considers that "nonparametric methods are in fact based on population parameters, and that these parameters should be estimated, with sample statistics and confidence limits, instead of following the traditional practice of calculating only P-values for the sample statistic" (Newson, 2006a, p. 309).

[293] Somers' D is an ordinal measure of association and was introduced by Somers (1962).

[294] The z-scores are calculated as $z = \frac{estimated\,value}{standard\,error}$ and enable a comparison of the effect size of different variables.

Table 5.6: Trend analysis of core performance

EXPLANATORY NOTE
The table shows the results of the trend analysis on the performance of the core in predicting the collective decisions. It contains the number of observations, the Somers' *D* coefficient, its z-score and 95% CI. The observations are separated by procedure and for being the first or second decision rule of a session. Statistically significant (two-tailed) at the 0.05 level ** and at the 0.01 level ***. The unit of observation are collective decisions.

Independent variable: NUMBER OF ROUND		Dependent variable: SELECTION OF THE CORE N	*D*	z-score	95% CI	
PROCEDURE	pooling	58	-0.04	-0.45	-0.21	0.13
	first decision rule	18	0.09	0.55	-0.25	0.44
	second decision rule	40	-0.11	-1.04	-0.32	0.10
	simultaneous delegation					
	delegation level	84	0.06	0.83	-0.8	0.20
	first decision rule	56	0.15	1.73	-0.02	0.33
	second decision rule	28	-0.28**	-2.43	-0.51	-0.06
	sequential delegation	48	0.14	1.42	-0.05	0.33
	first decision rule	30	-0.05	-0.36	-0.31	0.22
	second decision rule	18	0.43***	3.43	0.19	0.68
	all	148	0.45	0.91	-0.05	0.14
	first decision rule	76	0.05	0.62	-0.10	0.19
	second decision rule	72	-0.09	-1.21	-0.24	0.06

repeated play led to a better performance of the core. This is not surprising as the sequential decision-making facilitated the coordination of the two delegations by reducing the level of uncertainty. All SIM rounds as second decision rule were run after playing under POL before (cf. Sec. 4.4). The negative trend tells us that the selection of the credible core became less and less the longer the experiment continued. As the trend was discovered at the delegation level it is not artificially caused by aggregation. Rather, the complexity of SIM crowded out the amount of theory-compliant behavior.

This evaluation of the theoretical benchmark prediction made two aspects very clear: i) the overall performance of the core was around 40% which left some space for other explanatory factors and ii) there was no common trend over the rounds for the core performance. Due to the complexity of the task the majority of participants seemed not able or not willing to follow the theoretical benchmark. Yet, if the solution concept was not decisive, what were the crucial criteria for selecting an alternative? The next subsections look into this question.

5.4.2 Social welfare allocation and stability

Assessing the social welfare allocation I restricted my analysis to payoff tables which represent non-constant-sum games.[295] Tab. 5.7 displays the results. The trend analysis (cf. Tab. A.9) shows no significant changes for any procedure. Also, combining results across all procedures and looking only at decisions made when used as first or second decision rule made no difference.

[295] To look into what a group has earned when each option summed up to the same amount of points would make no sense.

Table 5.7: Social welfare allocation

EXPLANATORY NOTE
The table shows the social welfare allocation of the collectively selected alternative for the different procedures. The observations are separated by procedure and for being the first or second decision rule of a session. In addition, I differentiate under SIM for group and delegation level and under SEQ for first and second stage. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE	N	mean	SD	min	max
pooling	10	64.2	15.2	39.3	77.3
first decision rule	2	68.6	4.4	65.5	71.7
second decision rule	8	63.2	17.0	39.3	77.3
simultaneous delegation	8	55.0	7.6	45.9	66.4
first decision rule	6	55.7	8.8	45.9	66.4
second decision rule	2	52.6	0.9	52.0	53.3
delegation level	8	50.8	0.9	50.1	52.8
first decision rule	6	51.0	0.9	50.2	52.8
second decision rule	2	50.4	0.4	50.1	50.7
sequential delegation	8	63.0	12.8	41.5	87.2
first decision rule	5	66.5	12.3	56.7	87.2
second decision rule	3	57.1	13.6	41.5	66.2
delegation level	8	53.7	4.2	44.6	58.9
first stage	8	43.8	16.1	22.0	74.2
second stage reduced	8	68.4	27.4	15.0	100.0

Overall, subjects were able to obtain most points under POL. In particular compared to SIM the distinction is notable ($p = 0.096$). At the delegation level, SEQ significantly outperformed SIM ($p = 0.006$). Of course, for SIM I used probabilistic outcomes due to the selection of sets instead of alternatives by the participants. This results in values around 50% for both the maximum and minimum of points under SIM.

Within POL the differences between first and second decision rule are negligible. The same is true for SIM, while under SEQ the welfare allocation is higher as first decision rule ($p = 0.053$). Also evident is that under SEQ the second stage surpassed the first stage ($p = 0.093$). Please note that I restrict the evaluation to the actual remaining choice of the second stage delegation. The term "reduced" refers to the problem of choice when only one remaining set is left over from the first stage decision-making. Here, it becomes obvious that the second stage was able to secure on average 68.4 of the possible points. This represents the overall best performance.

So far I discussed the performance of the different procedures in relation to each other. Yet, it is worthwhile to also have an absolute point of comparison. Tab. 5.8 lists the social welfare allocation which would result if the group had in all decisions on tables with equilibrium in fact chosen the core alternative.

The comparison to the actually obtained results shows that under POL and SIM the groups achieved a higher welfare allocation as predicted by the core. Yet, only POL performed significantly better ($p = 0.009$). Please note, that any deviation from the core alternative which resulted in more welfare for the group cannot, by definition, have been a Pareto-improvement.[296] Under SEQ the group ended up with fewer points ($p = 0.063$).

[296]For every core alternative it holds that under the rationality assumption no individual or subgroup within the coalition supporting it can be better off by deserting the coalition (cf. Peleg and Sudhoelter, 2003).

Table 5.8: Social welfare allocation of core alternatives

EXPLANATORY NOTE
The table shows the social welfare allocation which would result if the group had always chosen the core alternative. Obviously it includes only the observations made with tables containing such equilibrium. Please note that the different procedures lead to different core alternatives. The data is separated according to decision rule.

PROCEDURE	N	mean	SD
pooling			
always core	9	30.9	20.9
actual result	10	62.0	33.5
simultaneous delegation			
always core	7	44.9	35.4
actual result	8	55.1	8.6
sequential delegation			
always core	8	76.3	13.6
actual result	8	63.2	24.2

A second measure that was mentioned in my hypotheses is the stability of collective decisions. I approximate stability by means of a dyadic comparison. This comparison analyzed the collective decisions and individual votes when subjects were exposed to the same situation, i.e., deciding on the same payoff table in the same role.[297] Although this is only an auxiliary measure, the comparison provided some evidence on how uniform or different the voting patterns occurred. Standard errors and CI were obtained using STATA's binomial CI calculator.[298] More precisely, despite the programs' default option of binomial exact intervals (cf. Clopper and Pearson, 1934) and the "nearly universal use" of the Wald interval (Brown et al., 2001, p. 115) I calculated "Wilson CI" (Wilson, 1927). Following Brown et al. (2001), in comparison this method is more powerful, parsimonious and less error prone. The results are shown in Tab. 5.9.

Table 5.9: Stability within a voting procedure

EXPLANATORY NOTE
The table shows the approximated stability of the decisions among the different procedures. The data is separated according to decision rule and differentiates between collective and individual decisions. The percentage of corresponding decisions is determined as dyadic comparison of identical outcomes in similar situations (i.e., the same payoff table) within procedures.

PROCEDURE	N	% of corresponding decisions	SE	95% CI	
pooling					
collective	66	19.7	4.9	11.9	30.8
individual	1878	24.3	1.0	22.4	26.3
simultaneous delegation					
collective	75	22.7	4.8	15.7	33.3
individual	2394	55.3	1.0	53.3	57.3
sequential delegation					
collective	48	37.5	7.0	25.2	51.6
individual	1140	53.5	1.0	50.6	56.4

In this table the number of observations does not refer to the vote of a subject or a collective decision. Instead it depicts the frequency of occurrence of a specific choice situation

[297] As discussed in Sec. 4.3.1, I restricted the analysis of the individual votes to the last ballot in each decision-making. Always, these final votes of individuals sum up to the collective outcome.
[298] Source: `http://www.stata.com/help.cgi?cii` (accessed February 25, 2013).

with which more than one subject, delegation or group was confronted. The number results as $N = \binom{n_{identical\,situations}}{2}$ given $n_{identical\,situations} \geq 2$.[299] The computations show that stability at the collective level is not high under any procedure. In relative terms, POL was as stable as SIM. SEQ outperformed POL ($p = 0.034$) and SIM ($p = 0.074$); this indicates that the lower level of complexity due to the introduced sequence facilitated the collective coordination. At the individual level the variance under POL was clearly the highest (SIM: $p = 0.000$; SEQ: $p = 0.000$). The delegation procedures no longer differ and reach at least a level of 50% corresponding decisions.

One interesting facet represents the comparison across procedures. Here, I looked into the question how the votes differ between subjects voting on the same payoff table but under various procedures. Tab. 5.10 shows that at the collective level the decisions vary strongly between all procedures. Looking at the individual level, the highest coherence was observed, as one might have expected, between SIM and SEQ. In both procedures ballots were held in the group of three and every individual can only influence one dimension of the decision-making (either column or row). Here, at least 50% of the individual votes matched. Interestingly, individual votes under POL corresponded in the same amount (around 40%) to SIM as well as to SEQ.[300]

Table 5.10: Stability across voting procedures

EXPLANATORY NOTE
The table shows the approximated stability of the decisions among the different procedures. The data is separated according to decision rule and differentiates between collective and individual decisions. The percentage of corresponding decisions is determined as dyadic comparison of identical outcomes in similar situations (i.e., the same payoff table) across procedures.

Comparison of individual votes among DIFFERENT PROCEDURES	N	% of corresponding decisions	SE	95% CI	
pooling and simultaneous delegation					
collective	105	4.8	2.1	2.1	10.7
individual	607	39.2	2.0	35.4	43.2
pooling and sequential delegation					
collective	88	10.2	3.2	5.5	18.3
individual	494	39.1	2.2	34.9	43.4
simultaneous and sequential delegation					
collective	89	12.4	3.5	7.0	20.8
individual	520	50.8	2.2	46.5	55.1

Of my formulated hypotheses two were concerned with social welfare allocation and stability. A better performance in both metrics was expected for the first stage compared to the second stage under SEQ (HYPOTHESIS 3) and for POL compared to SIM

[299] For example, three players in the same situation lead to three dyadic comparisons (A-B, A-C, B-C), four players in the same situation lead to six dyadic comparisons (A-B, A-C, A-D, B-C, B-D, C-D), five players in the same situation lead to ten dyadic comparisons (A-B, A-C, A-D, A-E, B-C, B-D, B-E, C-D, C-E, D-E), etc. Thus, the number of observations follows just the binomial coefficient $\binom{n}{k} = \frac{n!}{k!*(n-k)!}$ (Heinrich, 2006, p. 13 and 223) for $k = 2$ and $n \geq 2$. This relationship also applies to collective decisions.

[300] Here, I analyzed if the choice of a specific alternative under POL corresponded to the choice of the corresponding set under a delegation procedure. For example, if an individual under POL chooses alternative $\{C2\}$ a corresponding decision for a column delegation individual under SIM or SEQ is set $\{C\}$.

(HYPOTHESIS 4). HYPOTHESIS 3 is rejected; in fact, it is the other way around. The second stage outperforms the first stage which could not exploit moving first. HYPOTHESIS 4 is confirmed when looking at social welfare allocation as SIM clearly performed worse than POL (which is also the only procedure under which the welfare allocation significantly surpasses the theoretical baseline). However, this does not apply to stability. Here, only SEQ surpasses SIM. At the individual level POL received the lowest stability rating. This might result from different thresholds in reaching a collective decision. Under POL even with two diverging votes a majority could be reached. The next section picks up this argument when discussing the approval rate.

5.4.3 Distribution of wealth and approval rate

The distribution of welfare is nearly as important as its allocation in the first place. I looked into the income distribution of my results by using two empirical measures: the standard deviation as well as the Gini coefficient of an alternative. Compared to the previous assessment of social welfare allocation I could also include constant-sum tables into the analysis (which now considers all collected 221 collective outcomes). It holds for both metrics that high values reflect a large imbalance and low values represent a small amount of inequality. As equality might be considered as a generally desirable property, low scores represent a good performance. I start with the standard deviation as criterion for the relative inequality performance of a collective outcome. The results are shown in Tab. 5.11.

Table 5.11: Standard deviation

EXPLANATORY NOTE
The table shows the standard deviation of the collectively selected alternative for the different procedures. The observations are separated by procedure and for being the first or second decision rule of a session. In addition, I differentiate under SIM for group and delegation level and under SEQ for first and second stage. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE	N	mean	SD	min	max
pooling	10	56.0	11.8	32.5	72.0
first decision rule	2	46.1	19.2	32.5	59.7
second decision rule	8	58.5	9.6	41.1	72.0
simultaneous delegation	8	47.2	4.6	39.1	50.9
first decision rule	6	46.0	4.8	39.1	50.1
second decision rule	2	50.7	0.4	50.4	50.9
delegation level	8	46.0	2.6	43.3	51.4
first decision rule	6	45.9	3.0	43.3	51.4
second decision rule	2	46.3	1.6	45.2	47.4
sequential delegation	8	62.4	13.8	43.1	80.9
first decision rule	5	68.8	12.8	51.9	80.9
second decision rule	3	51.7	8.3	43.1	59.8
delegation level	8	46.5	7.9	30.9	55.7
first stage	8	47.7	8.2	34.3	62.2
second stage reduced	8	48.9	13.5	30.3	67.1

Contrasting first and second decision rule across all procedures provides no findings except under POL ($p = 0.046$). The trend analysis (cf. Tab. A.9) shows no results except for one; weak significance suggests that under SIM as first decision rule the inequality decreased over the rounds.

The performance of SIM is good in general as the the smallest imbalances were obtained under this procedure. Comparing the results of SIM and SEQ as first decision rule, the inequality is significantly lower under SIM ($p = 0.006$). Looking at the realized minimal and maximal results the tendency of SIM towards 50% outcomes was not as pronounced as with respect to social welfare allocation (cf. Tab. 5.7).

As an absolute point of comparison Tab. 5.12 presents the standard deviation which would have resulted if groups had always agreed in tables with equilibrium on selecting the core. The juxtaposition shows no difference for POL. Under SEQ ($p = 0.036$) and SIM the groups ended up with a higher inequality ($p = 0.012$).

Table 5.12: Standard deviation of core alternatives

EXPLANATORY NOTE
The table shows the standard deviation which would result if the group had always chosen the core alternative. Obviously it includes only the observations made with tables containing such equilibrium. Please note that the different procedures lead to different core alternatives. The data is separated according to decision rule.

PROCEDURE	N	mean	SD
pooling			
always core	9	61.2	22.8
actual result	10	56.0	11.8
simultaneous delegation			
always core	8	33.5	13.3
actual result	8	47.2	4.6
sequential delegation			
always core	8	48.9	13.5
actual result	8	62.4	13.8

I also used the Gini coefficient to assess the relative inequality performance. This is a common measure for the distribution of income within a society (cf. Haughton and Khandker, 2009). Since I do not intend to show duplicated results in detail, I moved them in Sec. A.13. Here, Tab. A.10 displays the results for the Gini coefficients under the different procedures. The table also shows the Gini coefficient which would have occurred if the group had always chosen the core alternative. Overall, the comparison of the Gini coefficient with the standard deviation increased the reliability of my analysis as the values confirmed the previous results. Firstly, the good performance under SIM.[301] Secondly, as when considering the standard deviation, the comparison with its counterpart, i.e., the actual obtained results in equilibrium tables, is not decisive for POL. Under SEQ ($p = 0.012$) and under SIM ($p = 0.012$) the results of the collective decision-making possessed a higher Gini coefficient than the core alternatives.

In my hypotheses the distribution of points was linked to the approval rate of a decision. How many votes each decision received is listed in Tab. 5.13. Please note that the value range of the approval rate is from 66.6 ($\frac{2}{3}$ majority) to 100 (unanimity).[302] Already at first

[301] At the group level SIM outperforms SEQ ($p = 0.100$).
[302] Under POL two different majorities are possible, i.e., 4 or 5 out of 6 players.

glance it is clear that majority decisions were dominant across all procedures and modifications. Especially under POL unanimous votes were very rare. The only significant difference is found between the same subjects that acted first under POL and subsequent under SIM ($p = 0.080$). Here, unanimity was still unlikely but occurred at least to a small extent under SIM. The trend analysis (cf. Tab. A.9) suggested more unanimity decision when SEQ was used as second decision rule.

Table 5.13: Approval rate

EXPLANATORY NOTE
The table shows the approval rate for the different procedures. The observations are separated by procedure and for being the first or second decision rule of a session. In addition, I differentiate under SEQ for first and second stage. As both delegation procedures vote in delegations (of three players) I show the approval rate at this level. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE	N	mean	SD	min	max
pooling	10	68.71	1.86	66.67	72.22
first decision rule	2	69.44	1.96	68.06	70.83
second decision rule	8	68.52	1.92	66.67	72.22
simultaneous delegation	8	70.47	3.19	68.06	77.78
first decision rule	6	70.25	3.71	68.1	77.78
second decision rule	2	71.12	1.06	70.37	71.88
sequential delegation	8	71.61	3.97	66.67	77.78
first decision rule	5	70.28	3.31	66.67	75.00
second decision rule	3	73.84	4.62	68.75	77.78
first stage	8	70.49	7.99	66.67	88.89
second stage	8	72.74	6.00	66.67	83.33

Two hypotheses were directed to the relative inequality distribution and approval rate. A lower inequality and a more encompassing agreement were anticipated for the first stage in relation to the second stage under SEQ (HYPOTHESIS 5) and for SIM compared with POL (HYPOTHESIS 6). HYPOTHESIS 5 is rejected. I could not observe more unanimity voting in the second stage and found no differences in distributional patterns across the two stages under SEQ. HYPOTHESIS 6 is only partially confirmed. For inequality the findings show that SIM outperformed SEQ (but not POL) at the delegation level. With respect to the approval rate it is undeniable that majority and not unanimity decisions dominated. I found in particular very little unanimity under POL.

5.4.4 Decision-making efficiency

A central aspect of decision-making procedures is their efficiency. How long does it take to reach a collective agreement? The answer to this question is shown in Tab. 5.14. Of course, with respect to the two delegation procedures the efficiency of reaching a decision in a delegation was important. Therefore, I also list these results. While the previous discussed effectiveness measures hardly showed a change over time this was different for efficiency. But, contrary to expectations, the number of necessary ballots increased over the progressing rounds (cf. trend analysis in Tab. A.9) and from first to second decision rule ($p = 0.000$) when looking at all procedures together.

Table 5.14: Decision-making efficiency

EXPLANATORY NOTE
The table shows the efficiency of collective decision-making for the different procedures. The observations are separated by procedure and for being the first or second decision rule of a session. In addition, I differentiate under SIM for group and delegation level and under SEQ for first and second stage. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE	N	mean	SD	min	max
pooling	10	5.25	1.16	3.25	6.9
first decision rule	2	6.08	1.18	5.25	6.92
second decision rule	8	5.05	1.13	3.25	6.5
simultaneous delegation	8	1.83	0.29	1.25	2.25
first decision rule	6	1.74	0.27	1.25	2.00
second decision rule	2	2.07	0.26	1.89	2.25
delegation level	8	1.43	0.16	1.13	1.69
first decision rule	6	1.38	0.14	1.13	1.50
second decision rule	2	1.59	0.12	1.50	1.69
sequential delegation	8	2.72	0.22	2.38	3.00
first decision rule	5	2.63	0.20	2.38	2.88
second decision rule	3	2.88	0.22	2.63	3.00
delegation level	8	1.36	0.12	0.19	1.50
first stage	8	1.62	0.23	1.25	2.00
second stage	8	1.10	0.09	1.00	1.25

Looking at the data, two aspects stand out under POL. Firstly, no group was able to agree in the first ballot. This corresponds well to the previously discussed low probability of agreeing in a one-shot ballot under this procedure. Secondly, while the minimum number of ballots necessary was 2, at one instance impressive 32 ballots were required before a majority agreed to vote for the same alternative.[303] Turning towards the statistical evaluation, the trend analysis found under POL a decrease of efficiency over the rounds and when used as second decision rule (cf. Tab. A.9).

Under both delegation procedures the efficiency measures were considerably better than under POL. No round needed more than seven ballots and most agreements were already reached after 3 ballots were hold. This fast decision-making was stable for first and second decision rule as well for the delegation level. Comparing the observations across procedures shows that SIM ($p = 0.053$) and SEQ ($p = 0.046$) clearly outperformed POL.

Due to the sequential process it is not possible for SEQ to reach consensus on a collective decision of the whole group in only one ballot. Even if both delegations agree in their respective first vote, two ballots are always necessary. In the first ballot alone the first stage delegation participated. Only if this delegation immediately reached an agreement could the first vote of the second stage be held next. Thus, the undistorted values could only be obtained at the delegation level. Here, the efficiency of SEQ was as high as for SIM. Furthermore, when used as first decision rule SIM showed a higher efficiency than as second one ($p = 0.088$) and the second stage of SEQ possessed the highest overall efficiency and was significantly faster than the first stage ($p = 0.012$).

Sec. 5.3.2 formulates two hypotheses which were concerned with the decision-making efficiency. A faster collective agreement was expected for the less complex problem of

[303]The case of 32 ballots may constitute a special outlier as the next highest number was 13 rounds.

the second stage in relation to the first stage under SEQ (HYPOTHESIS 7) and for the smaller electorate under SIM compared with POL (HYPOTHESIS 8). Both hypotheses are confirmed. In fact, both delegation procedures showed clear efficiency advantages over POL. Surprisingly, the efficiency decreased in the number of rounds (cf. Sec. A.12) and from first to second decision rule.

5.4.5 Robustness

This section serves as robustness check for the previous discussed analysis. It is intended to correct alleged findings which resulted not from behavioral patterns but artificial table compilations. The hypotheses 3, 5 and 7 addressed differences in the performance of first and second stage participants under SEQ. SIM mirrored the decision-making of SEQ except for the sequential voting. Due to the instead simultaneously hold ballots no differences between the two delegations were expected. Thus, the juxtaposition of the performance of these two delegations enables a determination of the robustness of the treatment effects. Tab. 5.15 shows the results for the two delegations under SIM.

Table 5.15: Robustness of treatment effects

EXPLANATORY NOTE
The table shows the social welfare allocation, standard distribution, Gini coefficient and decision-making efficiency of the two delegations under SIM. The designation of the groups is based on their responsibilities in reaching a collective decision, i.e., to vote for either a column or a row (cf. Sec. 4.2.6). The unit of observation are session averages of the collective decisions under SIM.

SIMULTANEOUS DELEGATION	N	mean	SD	min	max
SOCIAL WELFARE ALLOCATION					
column delegation	8	45.0	4.8	35.9	52.4
row delegation	8	56.7	3.9	53.1	64.9
STANDARD DEVIATION					
column delegation	8	55.6	7.3	40.3	62.3
row delegation	8	55.9	10.2	46.9	78.9
GINI COEFFICIENT					
column delegation	8	0.43	0.09	0.23	0.52
row delegation	8	0.41	0.04	0.36	0.46
DECISION-MAKING EFFICIENCY					
column delegation	8	1.58	0.32	1.00	2.00
row delegation	8	1.29	0.15	1.00	1.50

HYPOTHESIS 3 stated a first-mover advantage under SEQ in terms of social welfare allocation and stability. I rejected this hypothesis as I found that the second stage performed better. This needs to be corrected slightly as the row delegation under SIM, which mirrors the second stage under SEQ, also outperformed its counterpart ($p = 0.012$). The difference is smaller as under SEQ but still significant.

With respect to HYPOTHESIS 5 I found no differences in distributional patterns for the two delegations under SIM. Neither the robustness check for standard deviation, Gini coefficient or approval rate displays significant values.

Following HYPOTHESIS 7 I expected a higher decision-making efficiency for the second stage under SEQ. This was found in the analysis. Looking at SIM, I obtained a somewhat higher efficiency for the row delegation ($p = 0.076$) which mirrors the second stage under SEQ. Again, the difference is smaller than and not as significant as under SEQ. However, also these findings need to be taken into account.

5.5 Chapter summary

This chapter started the empirical analysis of my experiment. First, Sec. 5.1 proves the reliability of the experimental design. I analyzed my data with a method previously used by S&K (2010). As the design of the experiments is similar but not equal, the compilation showed similar but not identical findings. However, my results indicate the same patterns as found in previous work. This fortified the reliability of my most conventional treatment. Next, I set up a benchmark for the upcoming investigation of collective results. Firstly, I explain in Sec. 5.2 the derivation of the credible core for all decision procedures. Secondly, Sec. 5.3 is concerned with further aspects which might have driven the behavior of the experimental subjects. The section discusses the trade-off between reaching a decision in an appropriate amount of time and the quality of this settlement. A long-established assessment when it comes to collective decision-making (Buchanan and Tullock, 1962; Kaiser, 2007a). In accordance with this literature, I formulate hypotheses for my design. Those are investigated in Sec. 5.4 empirically.

Overall, the theoretical solution concept could explain approximately 40% of the observations. This left room for other factors which might have driven the behavior of the experimental subjects. It also reinforces my argument for the additional assessment criteria for collective decisions.

Looking at those criteria I found, unsurprisingly, that delegation increased efficiency in terms of decision-making speed and that subjects facing less complex tasks agreed faster. Interestingly this was not true for experience with the game. Weak evidence showed a slower decision-making over the progress of the experimental sessions. Contrasting this efficiency gain against the effectiveness of the procedures provided ambiguous insights.

The review of the existing theoretical literature disclosed two arguments on the effect of nonseparability (cf. Sec. 2.5). Most importantly, ARGUMENT 2 expected a first-mover advantage when a decision characterized by nonseparable preferences is taken separately and sequentially. For this no evidence is found. With respect to social welfare allocation and decision-making efficiency my results show that it is rather the other way around. In consideration to the robustness check (cf. Sec. 5.4.5) this argument must be maybe somewhat attenuated, but a confirmation of the hypothesis is far away.

The proof for ARGUMENT 3 which expected a sub-optimal outcome when a decision characterized by nonseparable preferences is taken separately and simultaneously is sufficient. The social welfare allocation was highest under the pooled decision mechanism. Thus, for policies characterized by nonseparable issue dimensions the delegation of competencies must be well coordinated to prevent sub-optimal behavior. But if the term of effectiveness is understood in a more comprehensive way the results become less clear. Under POL the outcomes possessed also the most uneven distribution of points and the lowest amount of stability in collective results. Yet, in terms of stability SIM performed worse when compared to SEQ. It holds across all procedures that different decision mechanisms lead to substantially different outcomes. This process dependency is an important insight.

These findings are not related to the approval rate. This rate stays nearly unchanged across procedures and over experimental rounds; majority outcomes clearly dominated the ballots in my experiment. This suggests that subjects used their tactical opportunities; they excluded supernumerary players, formed minimal winning coalitions, etc. At this level of analysis it is not possible to determine if the outvoted players are intractable blockers or just poor outcasts.[304]

In fact, the small number of decisions which was reached by unanimity is the starting point for the next chapter which extends the analysis to the individual level. Under all procedures and table characteristics nearly always an outvoted minority of players existed. If my analysis would stop at the collective level the votes of these subjects would be lost. The final outcome only reflects the votes (and preferences) of the winning majority. Therefore, I look into the individual voting behavior in Chap. 6 and determine the individual driving forces for the observed collective results.

[304] Chap. 7 looks in detail into this matter. Here, I discuss the results of a post-experiment survey which asked the subjects for their criteria when allocating their vote. One question in particular focused on which other players the participants had considered when making their decision.

6 The determinants of individual choices

In this chapter I look into the disaggregated results of my experiment. This enables me to examine whether and, if so, how nonseparability affects individual behavior. In particular, I discover the individual driving forces for the identified collective behavior. As individual intentions and collective results do not always go hand in hand, this facilitates a much more distinct analysis. Using individuals as observational units enables me to fade out the noisy patterns of collective majority decisions (Halfpenny and Taylor, 1973). In my experiment some individuals are outvoted and, thus, the collective decision does not represent their choices. As in first-past-the-post voting the collective outcome does not indicate dissenting voices. Absolute majority ballots are held in groups of six and three members in my laboratory experiment. This raises the question of whether the exclusion is extensive enough to cause such trouble. Its maximum is reached with two of six or one of three votes. Yet, Goeree and Holt (2001, p. 4) argued that models "that introduce (possibly small) amounts of noise into the decision-making process can produce predictions that are quite far from any Nash equilibrium".[305] In other words, even the smallest deviations can exert a significant influence when incorporated in an analytical setting.

The following sections take a look at all final individual votes the participants made in the experiment, i.e., the single votes which made up the collective decision. The analytical process is similar to the preceding chapter. At first, in Sec. 6.1 I discuss previous literature on behavioral patterns of individual decision-making. Sec. 6.2 links the distinctive features of my design, complexity and uncertainty, to individual voting behavior and the participant's considerations. I discuss the implications of individual voting considerations and derive the resulting expectations. Thereafter, Sec. 6.3 brings the behavioral patterns and voting considerations together for the empirical analysis. I present in detail the composition of the statistical concept and its variants. In Sec. 6.4 I contrast the performance of various model specifications to determine the best with regard to my observations by means of a model comparison. The units of observation for the statistical model are all final votes of each individual in the final ballot of an experimental round. Finally, Sec. 6.5 summarizes the findings.

The main empirical results of this chapter are also discussed in "Delegation, Uncertainty and Social Preferences in Majority Decisions?" (Fleig and Finke, 2013). This article uses

[305] Cf. Crawford et al. (2010, Chap. 2) for a more comprehensive overview of equilibrium concepts that include noise in form of, e.g., some type of error or cognitive limitations.

the conducted laboratory experiment to investigate the variation of individual choice behavior in response to changing voting rules. As the goal of this chapter is similar, the empirical content of the two contributions largely matches. Compared with the article, the discussion of behavioral patterns of individual decision-making (Sec. 6.1) and participants' tactical considerations (Sec. 6.2) are expanded. Furthermore, additional models including various co-variates are examined (Sec. 6.4.2).

6.1 The literature on behavioral patterns of individual decision-making

The analysis in the preceding chapter clearly demonstrated that the cooperative game theory equilibrium only partly explains the observed behavior. Its performance in predicting the outcome of the laboratory experiment left a lot of room for additional explanatory aspects. The core alternative is characterized as a unique cooperative solution under the assumption of rationally acting individuals. A self-interested single player or subgroup has no incentive to opt for another choice. However, as not all collective decisions match the prediction, the question is what other aspects drive the voting behavior.

It is far beyond the scope of this study to provide an all-encompassing insight into behavioral research and theories on human behavior. A good starting point for such an objective is Cooper et al. (2007).[306] This section gives an overview of the most prominent, and, in the context of my experiment, relevant behavioral patterns of individual decision-making; i.e., I will not discuss endowment effects or status quo bias. Both are well-established behavioral patterns, but my experimental design implements neither an initial endowment nor an SQ.[307]

MORE THAN JUST PURE SELF-INTEREST

The tenet of the purely self-interested Homo economicus (cf. Persky, 1995) has provided a fruitful basis for many insights in social science (Kirchgässner, 2008). However, it is widely accepted that "real" people behave differently, especially if social interaction is involved (Coleman and Ostrom, 2009). Looking around, we see people living together in families, sharing their goods in communities, donating to the needy, etc. Even when we look as far back as Ancient Greece, Aristotle taught us that people are by nature social beings (cf. zoon politikon, Höffe, 2005). During the last two decades, behavioral economists and an increasing number of experimental social scientists have begun to systematically

[306] There are also valuable contributions which focus on more specific aspects; Sheppard (1998) and Camerer (2003) looked into game-theoretic models to explain behavior, Provis (2000) incorporated ethics, Charness and Garoupa (2000) discussed reputation, Bazerman and Lewicki (1985) as well as Mitchell (1985) highlighted aspects of negotiations within organizations, etc.

[307] Endowment effect (cf. Plott and Zeiler, 2011; Thaler, 1980) and status quo bias (cf. Knetsch and Sinden, 1984; Fernandez and Rodrik, 1991) describe a similar behavior; "once a person comes to possess a good, she immediately values it more than before she possessed it" (Tversky and Kahneman, 1991, p. 1041).

question the assumption that individuals only selfishly maximize their own material interest (cf. Palfrey, 2009; Henrich et al., 2001). In particular, the researchers aimed to identify the forces driving individual and collective decision-making (e.g., Ehrhart et al., 2007; Hastie and Dawes, 2010). Often, the explanatory power of their models fitted to experimental data could be successfully increased by allowing for additional aspects in an individual's utility function (e.g., Engelmann and Strobel, 2004; Löwenstein et al., 1989).[308]

These empirical studies experienced, as I did, that the actual results "deviate systematically from the game-theoretic prediction based on self-interest. These deviations are naturally interpreted as evidence of social norms (what players expect and feel obliged to do) and social preferences (how players feel when others earn more or less money)" (Camerer and Fehr, 2004, p. 90).[309] In contrast, not everyone and everything is focused on the common good. Most prominently, Smith (1776) taught us that progress and prosperity originate from self-interest. Not every action which has positive external effects has a social motive. For example, a baker makes loaves of bread to sell them, not to improve social welfare. Recent research in behavioral economics, social psychology and neurosciences has begun to reconcile these two perspectives on human nature (e.g., Bräutigam, 2005; Buller, 2005). Departing from the idea of man as Homo economicus it has become widely accepted that human interaction is partly motivated by both kinds of preferences (cf. Henrich et al., 2001). That includes both caring (e.g., fair and trusting) as well as misgiving (e.g., competitive and spiteful) intentions (cf. Eisenkopf and Teyssier, 2010; Engelmann and Strobel, 2004).[310]

6.1.1 Other-regarding preferences

The inclusion of additional aspects, besides one's own payments when making decisions, implies a change of focus. Unlike self-interest, which can simply be read as the payoff a subject receives, the considerations now take on a relative perspective. A famous quote of Helson (1964) states that a human's "perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes. When we respond to attributes such as brightness, loudness, or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point".[311] While this concerns the general way of per-

[308]Ockenfels (2007, p. 3) pointed out that to be different from Homo economicus does not imply that people behave irrationally or chaotic. Humans follow their own rationality, which may differ from pure self-interest; but their behavior still is systematic, predictable and, thus, can be modeled.

[309]In their contribution Camerer and Fehr (2004) pointed out the important role of experimental methods in this field of research. Within an experiment it is possible to "carefully control players' strategies, information, and possible payoffs" (Camerer and Fehr, 2004, p. 90). This secures the necessary validity for identifying the behavioral aspects.

[310]Levitt and Dubner (2009, 2011) explain manifold and in detail that "people aren't 'good' or 'bad'. People are people, and they respond to incentives. They can nearly always be manipulated - for good or ill - if only you find the right levers" (Levitt and Dubner, 2011, p. 125).

[311]Kahneman and Tversky (1979, p. 227) illustrated this vividly as "an object at a given temperature may be experienced as hot or cold to the touch depending on the temperature to which one has adapted."

ception, the "same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another - depending on their current assets" (Kahneman and Tversky, 1979, p. 227). This insight is not only true for laboratory research or a specific environment. "Overwhelming evidence shows that humans are often more sensitive to how an outcome differs from some reference level than to the absolute level of the outcome itself" (Rabin, 1998, p. 4).

ABSOLUTE VS. RELATIVE GAINS

The experimental laboratory resembles a rather abstract environment for participants. They are admitted to a variety of roles and have to act accordingly. To judge and evaluate in this setup may be a challenging task. It is thus not surprising that subjects aim to assess their position with respect to a certain benchmark; while subjects try to evaluate their relative position, these "individuals desire to occupy a (subjectively) better position than their peers" (Ok and Kockesen, 2000, p. 533).[312] This attempt to be relatively better off is a continuous pattern found in many experiments (Kurzban and Houser, 2005). It is sometimes referred to as competitiveness (e.g., Crawford, 1985; Mitchell, 1985) and sometimes, more negatively, as envy (e.g., Kirchsteiger, 1994; Beckman et al., 2002) or spitefulness (e.g., Levine, 1997).

Using lottery and money-transfer games Kuziemko et al. (2011) identified a "last-place aversion". They found that the last-place player and the second-to-last-player (in terms of an income ranking) were willing to bear a high risk for the chance of moving up in rank in order to defend their position. The authors linked their results to the empirical question of why especially low-income individuals often oppose more redistributive policies. They concluded that people do not only passively compare their relative situation, but actively defend their interests. Ockenfels (1999, p. 2) argued that "pure altruism" is rare and that altruistic behavior is more likely to be driven by concerns for a specific relative position.

Looking at the importance of relative positions "one ubiquitous pattern stands out: [...] in a wide variety of domains, people are more averse to losses than they are attracted to same-sized gains" (Rabin, 1998, p. 5). The best known example of this is the work of Kahnemann and Tversky (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981).[313] These researchers showed that human choices depend on the outlook, or respectively the prospect, of the reference environment. It is crucial in which context or frame a person has to make a decision. Potential gains or potential losses are perceived differently and people try to prevent losses even at high costs.[314]

[312] This insight is a well-known fact which goes back at least to the "relative income hypothesis" of Veblen (1899).

[313] Cf. Heukelom (2007) for an overview on their work and the origin of behavioral economics.

[314] The above discussed and recognized patterns of behavior of endowment effect and status quo bias are perfectly in line with loss aversion.

In summary, humans look for reference levels to make evaluations. The more unusual or complex a decision situation, the more this occurs (Dufwenberg et al., 2008; Schmidt, 2010). My experiment did not introduce an SQ which could have served as a focal point. Thus, subjects with such (comparison) intentions were left with only one possible reference point, the payoffs of the other subjects. Lopomo and Ok (2001, p. 2) described such subjects as "interdependent" because their utility depends not only on their absolute level of earnings but also on the relative share of the total surplus. The question is with which intentions a participant assessed the other players' payoffs.

6.1.2 Social preferences

Camerer and Fehr (2004, p. 90) defined social preferences as "how players feel when others earn more or less money" than themselves. Many experimental investigations have found empirical evidence of such concerns (e.g., Charness and Rabin, 2002; Fischbacher et al., 2008). Those "are demonstrated to play a complex role in explaining cooperative behavior" (Frohlich and Oppenheimer, 1996, p. 117).[315] Most importantly, they could be identified in various settings and forms;[316] so Löwenstein et al. (1989, p. 426) reasoned that "people care about the outcomes of others. We sacrifice our own interests to help loved ones or harm adversaries." While many studies analyzed 2-player games (e.g., Bolton and Ockenfels, 2000; Fehr and Fischbacher, 2003) this also applies to collective decisions.[317] In particular, S&K (2010) explicitly designed their experiment to test for social preferences in majority decisions. Overall, Fehr and Fischbacher (2002) pointed out that the failure to model social preferences may misguide the researcher in answering fundamental economic questions.[318] Interestingly, Diermeier and Gailmard (2006) concluded that social preferences may exist but depend on the reservation value and decision-making process. This is another good reason to investigate my data for social preferences; I do not specify a reservation value and use different voting procedures.

SOCIAL WELFARE PREFERENCES

The most prominent models of social preferences implement a self-centered measure for fairness (Cox and Sadiraj, 2012) in which a subject's deviation from the payoff of others or the average payoff drives its utility. A closely related but different approach was put forward by Andreoni and Miller (2002) and Charness and Rabin (2002). These authors

[315] A common finding for collective settings is that in public good games a compulsory fee which is enforced on everyone if an initial subgroup reaches the threshold increases the chances of provision of the public good (cf. Dawes et al., 1986).

[316] Sec. A.14 gives an overview on experimental games which have frequently been used to measure social preferences.

[317] In general, the existing literature finds social preferences to be stronger in subjects who interact in small groups than in subjects who interact in large groups (Fehr and Schmidt, 1999). Yet, this statement is based on comparing very different sets of decision rules. On the one hand, we find highly structured 2-player games, whereas market competition dominates the study of larger groups (Schmidt, 2010).

[318] I will not deny that there are also more pessimistic voices. For a critical review on the empirical evidence for social preferences cf. List (2009).

advocate a model of social welfare preferences which is characterized by individuals who care about the size of the pie, i.e., the actors attach a positive weight toward the aggregated surplus off all participants.[319] Accordingly, they choose the alternative which maximizes their utility taking into account the sum of all actors' payoffs. Charness and Rabin (2002) carried out 32 experiments and found strong support for this model. However, S&K (2010), whose design is very similar to mine, found no evidence in favor in their experiment where "social welfare seems to be a poor explanation for behavior in majority decision making" (S&K, 2010, p. 675). I therefore apply a model of self-centered fairness. However, this model will also be able to reflect social welfare preferences if they prevail.

INEQUALITY AVERSION

In Sec. 5.1 I introduce the ERC model of fairness. Yet, in my individual level analysis I opted for a similar but different approach to incorporate social preferences. I used the model of inequality aversion proposed by Fehr and Schmidt (1999). It is formulated as stated in Eqn. 6.1. An individual's utility consists of three components: their own payoff y_i, the difference between their own and another's payoff if the individual gets less D_i^- and the difference between their own and another's payoff if the individual gets more D_i^+. Subsequently, I refer to D_i^+ as advantageous inequality and to D_i^- as disadvantageous inequality.[320] In the experiment the construction of the payoff tables constrains D_i^+ to the interval $[0; 100]$ and D_i^- to the interval $[0; 20]$.

$$U_i(y_i) = \alpha_i y_i - \beta_i D_i^- - \delta_i D_i^+ \tag{6.1}$$

, where $D_i^- = \frac{1}{n-1}\sum_{j \neq i} max\left\{y_j - y_i, 0\right\}$ and $D_i^+ = \frac{1}{n-1}\sum_{j \neq i} max\left\{y_i - y_j, 0\right\}$.

How does this model differ from ERC? Instead of comparing actor i's share to the globally equal distribution $(\frac{1}{n})$, the inequality aversion model suggests a pairwise comparison to all other actors $j \neq i$. For example, imagine a five-player committee decision in which player $i = 1$ is pivotal and able to choose between two alternatives with the payoff vector $\{A\} = \{\underline{10}, 20, 20, 0, 0\}$ and $\{B\} = \{\underline{10}, 10, 10, 10, 10\}$.[321] The ERC model cannot discriminate between the two alternatives. No matter how important social preferences are, player 1 gets a fair share of $\frac{1}{5}$ with either of the two alternatives. Thus, ERC ignores any inequality between other players (Engelmann and Strobel, 2000, p. 8). By contrast, in the inequality aversion model U_i depends on the exact values of δ_i and β_i. Although both models lead to similar predictions for two-player games, the predictions differ for multiplayer games (cf. Engelmann and Strobel, 2004).

In their theoretical discussion Fehr and Schmidt (1999, p. 822) assumed that i) $0 < \delta_i < \alpha_i$ and ii) $\delta_i < \beta_i$. The first assumption rules out two types of players: those who like being

[319] This applies only in the case of settings where side payments are not possible. Otherwise such behavior could also be led by strategic, instead of social, considerations.

[320] Goeree et al. (2002, p. 267) used the designations of "guilt" and "envy" to characterize the parameters.

[321] The underlined numbers indicate the payoff of player $i = 1$.

better off than others and those who place a higher importance on their relative gains than on their absolute gains. The second assumption states that a relative loss diminishes the utility more than an equally sized relative gain increases it, i.e., "losses resonate more than gains" (Rabin, 1998, p. 5). From a theoretical perspective, it is important to note that ERC and the inequality aversion concept have been developed for games in which a subject's preference for advantageous inequality does not alter equilibrium behavior, in particular the ultimatum game, the dictator game and public good games (Fehr and Schmidt, 1999, p. 851).[322] Accordingly, any theoretical discussion starts with the assumption that if social preferences affect individual behavior, they must reflect a concern for fairness and reciprocity.

This assumption does not hold true for plurality voting under majority rule as applied in my experiment. Here, the more general question is how far players are willing to forgo an increase in absolute payoffs to improve their own or other players' relative positions. To answer this question I could not use the ERC model as it does not differentiate between relative gains and losses. This was the main argument for using the inequality aversion model of Fehr and Schmidt (1999). However, the originally provided value ranges for the model parameters restrict the possibility to measure those aspects by focusing on socially desirable intentions. Yet, as the model provides both parameters nonetheless, I also used them to investigate possible envious preferences. I therefore did not implement in advance the defined range constraints but left them arbitrary.[323]

6.1.3 What makes people social?

Scientific research concerned with evolutionary dynamics and human nature indicates that mankind is far from homogeneous (Buller, 2005; Weibull, 1995).[324] This has led to an "increasing emphasis on the importance of individual differences in understanding and modeling behavior and dynamics in experimental games and decision problems" (Kurzban and Houser, 2005, p. 1803). In the context of social preferences this leads to the question of if and how individuals vary in their degree of cooperativeness. Looking at the interface between personality psychology and economics, Borghans et al. (2008) pointed out that in psychology it is a well-known fact that certain personality traits are more malleable than cognitive ability over the life cycle.

"Measuring the magnitude of the concern people have for others, sometimes called Social Value Orientation (SVO), has been an interest of many social scientists for decades" (Murphy et al., 2011, p. 771). This concept highlights the role of individual characteristics for

[322]Not surprisingly, games which showed the clearest evidence for the existence of social preference are dictator and ultimatum games (cf. Camerer and Thaler, 1955).

[323]Even Fehr and Schmidt (1999, p. 850) acknowledged that the assumption, according to which players who like to be better off than others do not exist, is "unsatisfactory".

[324]Questioning standard theorems of evolutionary psychology, Buller (2005) argued that human minds are not calibrated to one prehistoric base but continually adapt during both evolutionary time and individual lifetimes. This leads to the large number of varieties and generates a polymorphic population.

decision-making. Here, Bradler (2009, p. i) observed strong reactions "when the payoffs for both subjects can be perfectly equalized." She also found a decrease of cooperativeness when the other player was better off. Both aspects support the idea of inequality aversion.[325] Overall, she concludes "that a need for fairness and equality fundamentally influences individual decision-making processes" (Bradler, 2009, p. i).

Social psychology literature (e.g., SVO) suggests that people can be classified as competitors, cooperators, and individualists (Komorita and Parks, 1995). This research fits very well with work in experimental as well as behavioral economics that classifies people as "spiteful" (Herrmann and Orzen, 2008; Saijo and Nakamura, 1995), "selfishly payoff-maximizing" (Eckel and Grossman, 1998) or "altruistic" (Cason et al., 2004). Applying evolutionary simulations, Lomborg (1996) calculated a stable population of three types: cooperators, cautious cooperators and non-cooperators. This is exemplary for the literature in this field across the different disciplines. Most contributions end up with three different types: i) spiteful competitors trying to achieve better payoffs than others (cf. Sec. 6.1.1), ii) cooperators caring for others or aiming for a high welfare of the group (cf. Sec. 6.1.2), and iii) individualists who focus just on themselves.

Of course, once you classify people into these categories you are also interested in their proportion of appearance.[326] Here, Löwenstein et al. (1989) conducted an experiment in which subjects had to review different negotiations. They varied the dispute type between personal and business as well as between relationships designated as positive and negative. They categorize their subjects "according to their preferences toward advantageous inequality" (Löwenstein et al., 1989, p. 438). They identified saints, who preferred equality over inequality, regardless of the relationship, and loyalists, who preferred equality in positive but advantageous inequality in negative relationships. Ruthless competitors sought advantageous inequality under all conditions. The proportions of their sample amounted to 22% saints, 39% loyalists and 29% competitors. Andreoni and Miller (2002) tested the consistency of altruism in a modified version of the dictator game. They found that selfish preferences accounted for about half of the subjects, while about 33% of subjects divided equally between both players, and another 20% gave most tokens to the person with the highest redemption value. With public good games and an type-classification algorithm Kurzban and Houser (2005) identified 20% of their subjects as free-riders, 13% as cooperators and 63% as reciprocators. The authors pointed out that their distribution of types is similar to the findings of Fischbacher et al. (2001).

Fehr et al. (2001, 2002b) showed that these different types can be modeled accordingly by allowing for heterogeneous social preferences (HSP). In their work the explanatory power of the statistical model rests on the interplay of strictly egoistic as well as inequity-averse subjects. Traub et al. (2009) studied voting on redistribution. They found that in-

[325]On the one hand, subjects are willing to help others "even if it is costly to them" (Bradler, 2009, p. 17). Yet, on the other hand, envy causes a crowding out of benevolent behavior (Güth et al., 1982).

[326]Kurzban and Houser (2005) offered a comprehensive overview of literature in this field.

corporating HSP improves the predictions significantly compared to the standard model. Erlei (2004) and Dittrich and Ziegelmayer (2010) are further examples that build on the assumption of HSP. Dittrich and Ziegelmayer (2010) originally conducted bilateral gift exchange experiments to investigate the impact of loss aversion. After adjusting their theoretical predictions for social preferences, the authors aimed to identify the social intentions of individuals by endogenous types using a stochastic choice model. Their data can be explained best when allowing for at least four types of HSP. A promising approach is Erlei (2004), who combined social preferences, social-welfare preferences and reciprocity into one model of heterogeneous actors.[327]

With respect to my research it is important if these differences in people's fairness concerns matter for majority decisions. Hoechtl et al. (2012) argued that this depends on whether fair-minded voters are pivotal, which in turn is dependent on the distribution of voters and the design of the electoral system. Thus, the authors concluded that "fairness concerns matter when few fair-minded voters are sufficient to tip the balance in majority voting, but do not matter much when many are needed" (Hoechtl et al., 2012, p. 1416).

To summarize, the literature agrees that individuals vary in their degree of cooperation. The explanatory power of multiple models has been improved by accounting for this heterogeneity and allowing for deviations from rational choice. "However, there is little agreement concerning the best abstraction and the relative importance of the distinct psychological factors" (Ert et al., 2011, p. 258). In Sec. 6.3 I discuss co-variates whose impact on social considerations was assessed when estimating the statistical model.

6.1.4 Reciprocity

An important fact, which was also discovered in experimental behavioral research, is the dependency of behavior on accountability and context.[328] In other words, the role of personal involvement is of immense importance to an individual's actions (Branas-Garza et al., 2009). This is aptly illustrated by Dana et al. (2007) who have found that reducing accountability for actions in dictator games leads to significantly less generous behavior. The insights discovered on reciprocity "constitute a departure from neoclassical theory" (Cox and Sadiraj, 2012, p. 927). Here, depending on the prior actions of others, the intentions of subjects are altered. Deviating from "the conventional assumption that these preferences are stable [, ...] context-dependent preferences can capture the possibility that agents are motivated in part by reciprocity" (Sobel, 2005, p. 392).

With respect to the aforementioned social preferences it is important "to discriminate between behavior motivated by reciprocity and behavior motivated by nonreciprocal

[327] Perhaps the most unambiguous evidence for the usefulness of this approach is its ability to explain many behavioral anomalies (Goeree and Holt, 2001).

[328] Fehr and Schmidt (2000, p. 1) summarized that "in recent years experimental economists have gathered overwhelming evidence that systematically refutes the self-interest hypothesis and suggests that many people are strongly motivated by concerns for fairness and reciprocity."

other-regarding preferences" (Cox and Deck, 2005, p. 633). Altruism is very different from reciprocity. The first implies that an actor takes a costly action to increase the welfare of others irrespective of their prior behavior (Bradler, 2009). Reciprocity, on the other hand, means non-selfish action conditioned on the actions of others (Camerer and Fehr, 2004, p. 56). Such corresponding behavior is typically observed among principals and agents.[329] Most importantly, this might extend to vindictive actions as hostile acts provoke punishment intentions (Herne et al., 2012; Rabin, 1993).

With regard to the literature on reciprocity,[330] a wide variety of papers in economics (e.g., Andreoni et al., 2003; Bandiera et al., 2005), especially behavioral economics (e.g., Al-Ubaydli and Lee, 2012; Charness, 2004), and psychology (e.g., Baumeister et al., 2001) investigated the possible shape of a "reciprocity function". Although until now not all questions have been answered and a lot of research is still being conducted,[331] two main conclusions seem valid. Firstly, the shape of the "reciprocity function" is concave[332] (Al-Ubaydli and Lee, 2009; Bellemare and Kroger, 2007) and secondly, that negative reciprocity is clearly stronger than positive reciprocity (Baumeister et al., 2001; Offerman, 2002). Again, this corresponds well with the previously discussed human perspective that "losses resonate more than gains" (Rabin, 1998, p. 5).

Above all, the decisive aspect with respect to reciprocity is monitoring (Al-Ubaydli et al., 2010). Only when one observes the actions of others can one respond appropriately.[333] This holds true for individuals who are monitoring as well as being monitored (cf. Bandiera et al., 2005). Yet, it makes reciprocity less crucial for my experiment as my design explicitly introduces restricted monitoring capabilities. When deciding simultaneously, neither of the two delegations observes the vote of the other; and even when under SEQ the second stage is informed about the first stage decision before its vote, it is not possible to attribute the outcome to an individual but only to the complete first delegation. The applied majority bargaining makes the accountability of results towards one individual difficult and, thus, complicates a responsive reaction.[334]

6.1.5 Risk aversion

Up to now, I have determined that it is necessary to control for social preferences when assessing the individual votes of my experiment. A second, also important finding in

[329] As found, e.g., by Dittrich and Kocher (2011) when looking into a shirking game.

[330] Cf. Ockenfels (1999, Chap. V) for a comprehensive overview on reciprocal behavior.

[331] Cf. Cox and Sadiraj (2012) or Nicklisch and Wolff (2012) for an up-to-date overview.

[332] A concave reciprocity function implies that more trust, confidence or benevolence leads to more reciprocity, but at a diminishing rate.

[333] With respect to what 'respond appropriately' means, Rabin (1993, p. 1281) concluded that "people like to help people who are helping them, and to hurt those who are hurting them."

[334] Many findings of behavioral patterns in laboratory experiments presume the comparison of one's own payoff with others. Yet, such an evaluation becomes complicated when the setting is moved away from standard two-player games (Boyd and Richerson, 1988).

behavioral research is missing so far: risk aversion.[335] Most prominently, Kahneman and Tversky (1979) discovered that subjects reveal risk averse preferences over relative gains. A large amount of literature argues that risk averse actors chose alternatives which maximize their minimal payoff or minimize their maximal loss (e.g., Palacios-Huerta and Volij, 2008).[336] In ultimatum games the "fear of rejection" (Lopomo and Ok, 2001, p. 2) has been found to be responsible for a large part of the results (Camerer, 2003).[337] While game theory predicts that the proposer receives (nearly) all of the pie, this is rarely observed in empirical research (Bahry and Wilson, 2006). However, recent studies reported that the "lemon avoidance heuristic" (Ert and Erev, 2008) is also unusual.

Importantly, a subject's risk aversion is inherently a relative concept. Under conditions of uncertainty, subjects are not willing to risk being worse off than either their peers or the SQ.[338] However, one should not treat risk and inequality attitudes as equivalent. Experimental evidence suggests that these do not necessarily correspond (Traub et al., 2009). Thus, both should be assessed separately.

Yet, before I followed this approach I had to determine if risk aversion was relevant in my experimental design; but this assessment could be made manifestly. The level of uncertainty and its variation represents, besides complexity, the most prominent aspects and central treatment in my experiment. Thus, I could not leave such a commonly found aspect of human behavior out of my analysis.[339] In addition to social preferences I also controlled for risk aversion.

The standard approach of estimating risk aversion resonates as constant relative risk aversion (CRRA) where the utility of a subject for a given sum of money y is represented as in Eqn. 6.2 (Morton and Williams, 2010, p. 274).[340] Using this formulation a subject's revealed risk aversion parameter r is assumed to be constant and relative across the possible range of payments. Following Holt and Laury (2002) I assumed that for $r = 1$ the natural logarithm is employed.[341]

$$U_i\left(y\right) = \frac{y_i^{(1-r)}}{(1-r)} \tag{6.2}$$

[335]The above discussed framing effects and prospect theory (cf. Sec. 6.1.1) go hand in hand with the concept of risk aversion (Cox and Harrison, 2008; Heinemann, 2005) and the construct of concave utility functions (Chajewska and Koller, 1999; Gorman, 1968).

[336]Palacios-Huerta and Volij (2008) identified this strategy in strategic games with professional soccer players. The authors also clarified two designations, *minimax* and *maximin*; both pursue the same goal. Minimax is a decision rule that aims to minimize the possible maximum loss. Alternatively, maximin tries to maximize the possible minimum gain. Maximum loss and minimum gain refer both to the worst case scenario in which the subjects wants to be as successful as possible.

[337]Other important aspects are, of course, fairness and also partially confusion (Prasnikar, 1997).

[338]Schmidt (2010, p. 9) used the term "social risk aversion" to describe this tendency.

[339]Cf. Harrison and Rutstroem (2008) for a review on experimental evidence on risk aversion.

[340]Cf. Eeckhoudt et al. (2005) for a comprehensive overview on how to recognize, quantify, analyze and incorporate risk into decision-making processes. The volume focuses on economic and financial decisions as, e.g., portfolio choices.

[341]Holt and Laury (2002) also pointed out that the division by $(1-r)$ is necessary for increasing the utility when it holds that $r > 1$.

When estimating the CRRA parameter, $r < 0$ indicates a risk seeking individual, $r = 0$ a risk neutral person and $r > 0$ a risk averse character. Although I estimated the CRRA of my subjects, I am aware of the difficulty of generalizing it. Berg et al. (2005, p. 4211) pointed out "that researchers must be extremely careful in extrapolating a person's or group of persons', risk preferences from one institution to another. Without appropriate benchmarks on the preferences of individuals, researchers can mistake changes in behavior caused by risk preferences for change in behavior caused by other stimuli such as information or rule changes." The inclusion of risk aversion did not serve the purpose of comparing its value to other studies but rather of accounting for a qualitative pattern in subjects' behavior (cf. Sec. 4.3.3). In other words, it corresponds to the inclusion of a control variable which checks whether the results are robust for risk averse preferences.

6.2 Sincere and sophisticated voting

In economics, the market is the single most important mechanism for collective choice. Here, Smith (1962, 1964) clarified that social preferences are irrelevant in perfectly transparent markets, when separable preferences are assumed, because they do not affect individual equilibrium strategies. Accordingly, Dufwenberg et al. (2008) investigated whether agents with altruistic preferences differ from selfish agents in perfectly competitive markets. They defined a separability condition on the basis of whether agents' preferences can be represented by a weighted sum of internal utility functions. If the condition holds, "agents who care directly about the welfare and opportunities of others cannot be distinguished from selfish agents in market settings" (Dufwenberg et al., 2008, p. 631). In other words, it is impossible for any subject to enforce a fair or equal outcome if there is competition. By insisting on their fair share, the players only hurt themselves; they cannot prevent the other market participants from trading. Thus, a market in which agents have social preferences is observationally equivalent to a market in which each agent only cares about themselves (Schmidt, 2010, p. 6).

This changes once we move our perspective away from the perfect "coordination through the invisible hand of the price mechanism" (Larsson, 1993, p. 87). In the realm of politics, majority decisions replace markets as the single most important choice mechanism (Erlenmaier and Gersbach, 2001, p. 2).[342] Of course, most democracies do not actually use simple majority voting in their legislative processes but rather stabilize them by a system of checks and balances, the division of power and representative elements (Beitz, 1990, Chap. 4, p. 60). These systems allow for a certain degree of crossover between voting and trading, but standard plurality voting does not; here, votes are cast simultaneously. This results in complex and uncertain environments.[343]

[342] Majority voting constitutes the most widely spread decision rule for democratic government. Any democracy is commonly associated with political equality and majority rule (Saunders, 2010a,b).

[343] In economics, these aspects are summed up as nontransparent markets which cause risk aversion (Levy, 2006; Pratt, 1964).

With a view to general elections, the literature considers two types of voters or, respectively, voter behavior (cf. Clinton and Meirowitz, 2004; Herrmann, 2012).[344] At one extreme are those who give up any tactical considerations right away and decide to vote sincerely. Such a voter neither considers another subject's choice nor the logic of the collective decision problem; they are solely focused on their own payoffs. For the sincere type, discriminating between social preferences and self-interest is easily accomplished as long as the importance of social preferences is strong enough to influence individuals' choices. At the other extreme are those voters who make an effort not to waste their vote on a minority position and who therefore behave strategically (cf. Kramer, 1972).[345] Assessing the prevalence and realism of this voting type, Riker (1982a, p. 169) claimed that such "strategic voting is an ineradicable possibility in all voting systems [and] almost always present in legislatures". Costa-Gomes et al. (2001, p. 1193) defined this sophistication as "the extent to which behavior in games reflects attempts to predict others' decision, taking their incentives into account".[346]

Yet, even with perfect information on their own and others' preferences, voters will find it difficult to judge the consequences of their vote choice for two reasons. First, even if there is a unique equilibrium strategy, it may be difficult to identify. Second, others have the same problem, which only reinforces the voter's struggle with their own strategy. This uncertainty increases with the number of alternatives put to vote, the size of the electorate, and the possible arrangements of preferences among voters (Ordeshook and Palfrey, 1988).[347] Therefore it becomes difficult to infer preferences from given votes.

In my experiment the participants face a complex and uncertain environment. When investigating behavior displayed under such a challenging task, the key problem is to separate the effect of social preferences from subjects' cognitive short comings and human error (Goeree et al., 2002). The additional explanatory value of, e.g., incorporating fairness and extending the utility function (cf. Eqn. 6.1) depends on subjects' ability to calculate their resulting advantageous and disadvantageous inequality, which decreases with rising complexity.[348]

One solution to this problem is an experimental design which aims at controlling for social preferences. For example, in Johnson et al. (2002) other-regarding preferences of subjects were turned off by introducing robots as bargaining partners. Jacquemet and

[344] Cf. Merrill and Grofman (1999, part I) for an introduction to (spatial) models of voter behavior.

[345] Cf. Costa-Gomes et al. (2001) for an extensive overview on literature which studied strategic behavior and Herrmann (2012, p. 64) for contributions which have demonstrated strategic voting empirically.

[346] Another incisive description was given by Kramer (1972, p. 170) who clarified that for a sophisticated strategy "it is necessary for a voter, to make the best use of his vote, to attempt to predict the contingency likely to arise: that is to say, how the others are likely to vote."

[347] Ordeshook and Palfrey (1988, p. 441) pointed out that the distinction for sincere and strategic voting is not limited to elections but "represents an important contribution to our understanding of committees, of institutions, and of the opportunities to manipulate outcomes by the manipulation of institutions."

[348] Moreover, subjects now have to form beliefs regarding the relative emphasis other players place on social preferences as compared to their self-interest (Andersson et al., 2012). Obviously, the latter argument only concerns sophisticated types who try to anticipate other players' strategies and actions.

Zylbersztejn (2011) neutralized the effect of relative comparisons by altering the payoff structure towards more symmetry between their players. Another solution to this methodological problem is offered by Diermeier and Gailmard (2006) who imposed a bargaining protocol including a random assignment of roles and role specific reservation values. Subjects are then aware of their chances for either role as well as their actual role at the moment they make their particular choice. In summary, in situations which preclude either sophisticated voting or social motivation the model parameters of the other can be directly estimated. This does not hold true for plurality voting under majority rule, as in my design, where a subject's cognitive capacity is decisive for making the necessary (other-regarding) computations.

BOUNDED RATIONALITY

The distinction between sincere and sophisticated voting is supported by a further argument: the cognitive capacity of every individual is limited. This insight and the designation as bounded rationality go back to Simon (1957) who raised the questions of "how do human beings reason when the conditions for rationality postulated by the model of neoclassical economics theory are not met?" (Gigerenzer, 2008, p. 124). The subsequent research program argued that humans learn cognitive shortcuts in everyday exercises (Gigerenzer, 2008). Such heuristics guide their behavior with the result of more or less acceptable outcomes. Yet, they definitely prevent lavish considerations (Kahneman, 2011) by turning to "decision rules that reduces the computational burden to feasible levels" (El-Gamal and Grether, 1995, p. 1137). The question is what heuristics people use and in what environment they work or fail (Gigerenzer and Selten, 2001). Current research indicates that "a) individuals and organizations often rely on simple heuristics in an adaptive way, and b) ignoring part of the information can lead to more accurate judgments than weighting and adding all information, for instance for low predictability and small samples" (Gigerenzer and Gaissmaier, 2011, p. 451).

In addition to the confinement of humans' cognitive capacity in general, a laboratory experiment may restrict it further. Acting in experimental settings can be an awkward task depending on the game played. In my setup, subjects had to cope with four different decision situations, which introduced different levels of complexity and uncertainty (cf. Tab. 4.2). Taking all together, under such circumstances an actor often cannot undertake all required computations as its rationality is bound by cognitive limitations.

What does this mean for sincere and sophisticated voting in my experiment? The higher the level of complexity or uncertainty, the more sufficient simple approaches will be to understand the observed behavior. Contrary to this, any sophisticated equilibrium strategy will diminish because "uncertainty constrains the statistical relationship between their strategy choices players can bring about" (Crawford and Haller, 1990, p. 571). Thus, rational predictions were expected to explain more voting behavior under POL than under SIM as well as under the second stage of SEQ in comparison to the first one.

Yet, if basic patterns become more important, which will that be? Many heuristics assume a pre-existing social environment such as peer groups or a successful role model whose behavior will be imitated (cf. Kahneman et al., 1982). In the controlled environment of the laboratory this is not feasible. Also widespread are actors who choose the alternative promising the highest payoff; in other words, just "take the best" (Gigerenzer and Goldstein, 1999; Gräfe and Armstrong, 2010). Yet, in iterated games this choice will then be conditioned upon each alternative's cue validity (Rosch and Lloyd, 1978), i.e., an alternatives chance for success in a majority vote. In order to be selected an alternative needed at least $\frac{2}{3}$ of all participants voting for it.

WINNING COALITION

In Sec. 6.1.1 I discuss humans' inherent compulsion for comparison. When talking about the general principle, the aspect of group size seems unimportant, i.e., people compare themselves with one or several others all the time (e.g., Lopomo and Ok, 2001; Azar, 2010). Yet, when looking at the exact strength of this influence, the size of the reference group may well play a role. Levine (1997) examined results of ultimatum and centipede experiments with respect to altruism and spitefulness. He aimed to access the dispersion of these characteristics in the sample by varying the number of participants. While altruism was independent of the number of participants, he found less spitefulness in larger groups. Levine (1997, p. 614) argued that "one explanation of spite is that it is really competitiveness, that is, the desire to outdo opponents. In this case, it is not the total utility of opponents that matters, but some measure of their average or maximum utility." Thus, even if it seems reasonable to expect subjects to evaluate their payoffs in comparison to other players' payoffs, it is not clear how subjects perceive and evaluate these other payoffs.

Investigating centipede games, Tremewan et al. (2012) identified differences of experimental subjects with respect to ingroups and outgroups.[349] In my bargaining game under majority rule the collective agreements were reached within a minimal winning coalition between 80-96% of all decisions (cf. Sec. 5.4). The majority rule splits the subjects into an ingroup, i.e., the winning coalition, and an outgroup, i.e., the outvoted subjects. To belong to this winning coalition is an understandable desire for every subject as it has the advantage of being able to co-decide which alternative gets selected. Outvoted players have no influence on the collective result. This puts the subjects into a strategic dilemma. From the above we know that the relative position is important (Sec. 6.1.1), so subjects avoid low payoffs as others would be relatively better off.[350] On the other hand, if they aim too high, they might be outvoted by the other subjects and would probably end up with a very low turnout. Considering the previously discussed risk aversion (Sec. 6.1.5)

[349] Focusing on level-k estimations Tremewan et al. (2012) discovered a difference in beliefs about other players and emphasized the role of uncertainty. Participants behaved as if they could completely predict the actions of ingroup members but only partly the behavior of outgroup members.

[350] In non-constant-sum tables it is possible that a low payoff still goes along with being better off than others. But this is only possible in alternatives with a very low overall sum and, thus, low cue validity.

subjects were expected to try hard not to be left out and to keep someone beneath them. In their ultimatum experiment Güth et al. (1982) found that proposers just maximized given their fear of rejection. I anticipated that these considerations would also take place: subjects had to take into account the voting threshold when deciding on the alternatives. Therefore, I considered it likely that only alternatives receive votes which offer at least four players an appealing amount of points.[351] In my statistical model I controlled for this by implementing a specific winning coalition dummy in the strategic voting component.

6.3 Statistical model

Sec. 6.1 and Sec. 6.2 discussed well-known and verified patterns of behavior discovered in laboratory experiments. To assess their relevance for my design, I incorporated them into different model specifications and compared their relative performance in explaining the experimental results. Thus, I explored different utility functions of which each includes specific parameters to uncover the pattern that best matches my observations. This is similar to the approach of Ert et al. (2011) and their choice prediction competition for models of social preferences and to Smirnov (2009) who conducted a voting experiment and compared the explanatory power of random, sincere, strategic and risk averse models. Also, Goeree et al. (2002, p. 265) compared "likelihoods" of variant models to determine the adequacy of behavioral assumptions on altruism.

The underlying idea of the model to be estimated can be summarized as follows. When investigating how individuals decide when confronted with a discrete choice problem, the simplest form of model is represented by a standard conditional logit model. Here, an individual compares their payoffs across all alternatives to reach a decision. Correspondingly, when looking at individual behavior within collective decision-making, the most basic model takes into account only each player's own gains. Yet, subjects who also consider the payoffs for other players and act accordingly strategically can be more successful. Therefore, the payoffs of single alternatives have to be weighted with respect to their probability of constituting the collective choice.[352] Following my theoretical solution concept, the weight corresponds to the probability of an alternative to represent the core.[353] More specifically, the weight is determined as the probability of an alternative j to survive against all other alternatives in a pairwise comparison across all possible winning coalitions. In other words, if there exists just one other alternative k which at least one specific player constellation with a sufficient number of players prefers over j, the

[351] Under both delegation procedures two sub-committees have to decide with at least two out of three votes. Thus, under all procedures an alternative must have received at least four votes to be selected.

[352] For example, consider a collective choice of six players and an alternative like $\{88, 1, 1, 1, 1, 1\}$ with five low and just one high number of points. While this option might be tempting for player 1 it is certainly not for the other players. Therefore, strategic considerations assign only a low probability of being the collective choice to this alternative.

[353] This approach is different from the one I applied in Chap. 5 where I used a deterministic definition. In the following, the core alternative is determined probabilistically.

not-preferred alternative is not the core because the core does not lose in any pairwise comparison. This transformation follows the idea of QRE which adjusted the basic logit model for players' errors (McKelvey and Palfrey, 1995, 1998). Yet, it is unrealistic to expect that all participants behave that way. Therefore, a second transformation allows for sincere and sophisticated behavior (Sec. 6.2). This mixture stage has the advantageous properties that one does not need to know the distribution of behavior[354] and that the model allows for observed as well as unobserved heterogeneity.

Next, Sec. 6.3.1 explains the implementation of the behavioral parameters in a utility function. The operationalization of the random utility model is shown in Sec. 6.3.2. Here, I derive step-by-step the probabilistic individual choice function for sincere and sophisticated voting. Next, Sec. 6.3.3 introduces the mixture stage of the model and discusses various co-variates which might have influenced subjects' social or strategic considerations. Finally, Sec. 6.3.4 summarizes the different specifications for the empirical analysis.

6.3.1 The robustness of self-interest

This section focuses on the operationalization of the previously discussed behavioral patterns. The aspects considered are self-interest, social preferences and risk aversion. Following Bolton and Ockenfels (2000, p. 171), the extended (expected) utility function might be designated "motivation function because it emphasizes [...] a statement about the objectives that motivate behavior during the experiment." The function also contains a parameter of random shock which depicts a subject's errors due to a player's mistakes, cognitive incapacities, etc.[355] This follows Glasgow et al. (2012), who criticized the frequent use of conditional logit models when analyzing government choices. The authors argued instead in favor of a mixed logit extension which includes random coefficients to allow for unobserved heterogeneity.[356]

The utility level U_{ij} of a purely self-interested actor i is shown in Eqn. 6.3. It is solely given by their payment y_{ij} when alternative j is selected.

$$U_{ij}(y_{ij}) = y_{ij} \tag{6.3}$$

[354] When estimating the model I did not know whether or to which extent any player behaved strategically.

[355] The interpretation of such shocks differs between disciplines in social sciences. Psychology focuses on why a subject's behavior becomes more and more ambiguous when stimuli get weaker (Luce, 1959). Here, "stochastic elements represent intra-personal variations in utility levels, or perception errors, which may cause a subject to choose differently when faced with the same stimuli. In econometrics literature [...], stochastic choice models are usually applied to cross sectional data, containing decisions of many individuals. The stochastic elements are interpreted as inter-personal variation, or heterogeneity, in preferences. Experimental game theorists are mostly agnostic about the interpretation of the stochastic elements. In the laboratory, noise may be due to distractions, perception biases, miscalculations, or due to heterogeneous preference shocks such as feelings of envy, spite, or altruism. Regardless of the interpretation of the noise, the effect can be particularly important in an interactive context where players' payoffs are sensitive to others' decisions" (Goeree et al., 2005, p. 363).

[356] This argument also relaxes the independence of irrelevant alternatives (IIA) assumption.

Next, I looked for the robustness of this baseline assumption against random shocks and a player's concern for fairness. Thus, I was interested in the size of the parameters necessary to change the self-interest prediction. As I had no ex-ante knowledge as to whether a subject would vote sincerely (i.e., without consideration of other subjects' choices) or sophisticatedly (i.e., with a solution concept for the collective choice problem in mind), the operationalization had to accommodate both types.

I implemented the random shock ε_i as privately observed, mean zero random disturbances.[357] Furthermore, the use of independent and identically distributed (iid) extreme value error terms followed common logistic mixture model parameterization (Azaiez, 2010, p. 7). For sincere voting types the disturbance was operationalized by the relative value of the best as compared to the second best alternative. With respect to sophisticated voting types, I used the core as a theoretical solution concept. Thus, the error captured the necessary change for altering the core prediction. The inclusion of the random shocks leads to the utility function shown in Eqn. 6.4.

$$U_{ij}(y_{ij}) = y_{ij} + \varepsilon_i \tag{6.4}$$

When looking at concerns for fairness, I differentiated between sincere and sophisticated voting in the same manner. For sincere types I measured fairness by the relative difference b_i between the alternative with the highest score of the subject compared to the average and, thus, (theoretical) fair share of $\frac{1}{n}$ (cf. Eqn. 5.2).[358] As before when considering random shocks, this operationalization does not take into account the payoffs or strategies of other players. For sophisticated voting types I used again the core as benchmark. Here, I assessed the importance of fairness defined by inequality aversion (Fehr and Schmidt, 1999) as shown in Eqn. 6.5. In order to overcome the two-dimensionality of the concept, I identified the combined level of $(\delta_i + \beta_i)$ necessary to change the core prediction.

$$U_{ij}(y_{ij}) = \alpha_i y_{ij} - \beta_i D_{ij}^{-} - \delta_i D_{ij}^{+} + \varepsilon_i \tag{6.5}$$

In addition to random shocks and fairness I also investigated the impact of subjects' risk attitudes as the experimental design provided decision situations at different levels of uncertainty. Here, I followed the common approach of CRRA and transformed the utility function as discussed in Eqn. 6.2. Risk aversion extends to a subject's complete utility function. Thus, it influences an individual's payment as well as their social preferences,

[357] This corresponds to the original specification of the QRE framework (McKelvey and Palfrey, 1995).

[358] I used the ERC concept of fairness to account for social preferences of sincere voting types. Such subjects do not consider other subjects' choices or payoffs; but that is not necessary as ERC compares an actor's share only to the globally equal distribution ($\frac{1}{n}$).

which implies that the actor should dislike taking risks that are not taken by their reference group.[359] In this way the degree of risk aversion r transforms Eqn. 6.5 into Eqn. 6.6.

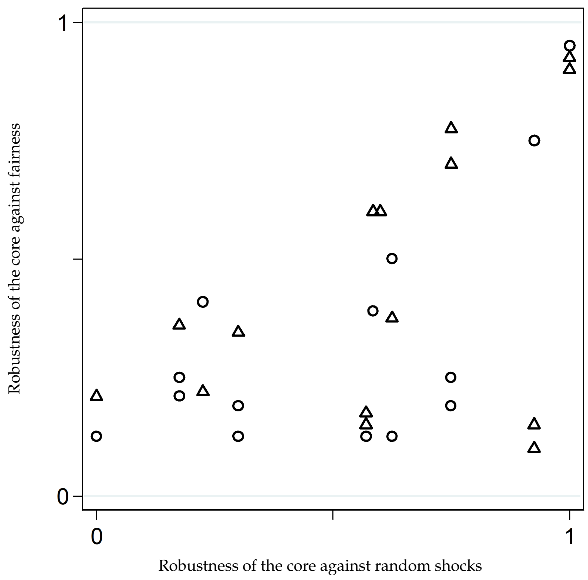
$$U_{ij}\left(y_{ij}\right) = \frac{\left(\alpha_i y_{ij} - \beta_i D_{ij}^{-} - \delta_i D_{ij}^{+} + \varepsilon_i\right)^{(1-r)}}{(1-r)} \tag{6.6}$$

ROBUSTNESS OF THE CORE

The payoff tables possess different degrees of robustness of the core against modifying the assumption of pure self-interest. This enabled me to estimate the respective parameters based on individual votes. Yet, before turning to the model implementation I look into the correlation between the different aspects. Fig. 6.1 reveals that the robustness of the core against random shocks and inequality aversion are correlated. The figure shows also fairness conceptualized along the ERC model (Bolton and Ockenfels, 2000, cf. Sec. 5.1). It is obvious that the correlation across all three characteristics is high.

Figure 6.1: Robustness of the core alternative

EXPLANATORY NOTE
The graph illustrates the correlation between payoff tables' robustness of the core against random shocks and fairness considerations. I use both fairness conceptualizations: inequality aversion (Fehr and Schmidt, 1999) is indicated as △ and ERC (Bolton and Ockenfels, 2000) is indicated as ○. The values are standardized and in the case of the inequality aversion concept the combined amount of both parameters is shown.



Due to this high correlation it was almost impossible to isolate the effect of any particular type of disturbance by analyzing purely collective choices.[360] This would bring about a high risk of confusing the effect of subjects' cognitive incapacities (e.g., random error) for

[359] Schmidt (2010) attributed this idea to explain herding strategies, where subjects spend their money on identical portfolios (e.g., houses) despite a known risk that prices are likely to drop.

[360] This is not unique to my experiment. A similar correlation can be found, e.g., in the payoff tables constructed by S&K (2010, p. 681-684). Their payoffs show a clear interrelation between the robustness against random shocks and social preferences.

social preferences or even strategic considerations. Fortunately, the prospects of identifying social preferences are less gloomy when analyzing individual choice behavior. This reinforced my decision to contrast collective and individual behavior. At the individual level, the greatest challenge is the limited ex-ante knowledge about a subject's level of sophistication. An appropriate statistical model therefore had to incorporate both sincere and sophisticated voting types.

6.3.2 Random utility model

Overall, my experiment represents a relatively unstructured voting game. In such a setting it is difficult to separate the effect of subjects' cognitive incapacities from behavioral parameters. Thus, I estimated random utility models to account for random disturbances in an individual's utility functions as, e.g., misjudgments and errors.[361] Following the approach of Goeree et al. (2002, p. 262) this was implemented by a probabilistic choice function which was supplemented by a theoretical equilibrium concept. This allowed for mixing sincere and sophisticated types as well as for estimating utility (or motivation) parameters in a combined model.[362] Following Harrison (2008, p. 2) this "joint estimation is essential for inference about the model as a whole". The optimization was accomplished through conditional logistic regression.

Most influential for the choice of this model was the contribution of Goeree et al. (2002). They played one-shot public goods games and looked for differences in internal returns for the subjects themselves and external returns to their co-players. The authors specified a logit equilibrium model in which choice was stochastic to obtain maximum likelihood estimations. In accordance with my model, the utility function of the subjects contained not only self-interest but also altruism parameters. Palfrey and Prisbrey (1997) used a similar approach when investigating warm-glow altruism in voluntary contributions experiments.[363] To assess the actions of their subjects the authors estimated response functions to treatments at the aggregated and the individual level using probit models.

SINCERE VOTING

I start by formulating the probabilistic choice of a sincere voting type. Such a subject is focused on their own payoffs and pays no attention to another subjects' (probable) choice. Thus, for sincere voting types their choice model is identical to the standard conditional

[361] Random utility models go back to the work of McFadden (1974, 1976, 1982). The approach has been extended to strategic settings several times (Signorino, 1999, p. 282), among others by McKelvey and Palfrey (1995, 1996, 1998) for their well-known QRE concept. For an up-to-date assessment on the empirical content and limitations of the concept cf. Haile et al. (2008).

[362] Signorino (1999) employed statistical strategic discrete choice models to a simple crisis interaction setup. Most important, he demonstrated how to directly incorporate the theorized strategic interaction into the discrete choice models.

[363] "Warm-glow preferences mean that the act of contributing, independent of how much it increases group payoffs, increases a subject's utility by a fixed amount" (Palfrey and Prisbrey, 1997, p. 830).

logistic regression.[364] This is shown in Eqn. 6.7 where p_{ij} is the probability with which an individual actor i votes for an alternative j. The probability results from the ratio of the, by individual's cognitive capacity λ weighted, utility U_{ij} of alternative j to the aggregate benefit of all nine alternatives $k = 1, 2, \ldots, 9$ (each also weighted by λ).[365] Therefore, every alternative possesses a positive probability to be chosen by a subject ($p_{ij} > 0$ for $j = 1, 2, \ldots, 9$). Obviously, "the choice probabilities are proportional to an exponential function of the expected payoffs" (Goeree et al., 2002, p. 263). Thus, "non-optimal choices can occur, but the probability of this is inversely related to their cost" (Goeree et al., 2002, p. 262). In terms of my experiment, the lower the payoff of an alternative j compared to the highest payoff of the subject in another alternative $k \neq j$, the smaller the probability of j to be chosen by the subject.

$$p_{ij} = \frac{exp\left(\lambda U_{ij}\right)}{\sum_{k=1}^{9} exp\left(\lambda U_{ik}\right)} \tag{6.7}$$

In situations where all subjects vote sincerely the model is identical to the standard conditional logistic regression, also at the group level. Thus, the probability of the collective outcome is simply the likelihood by which at least four players choose the same alternative coincidentally.

The additive utility function U_{ij} may be composed of various components. For example, absolute payoff or inequality preference and the respective parameters α, ε, and b which measure their influence. Importantly, these parameters and λ are not actor-specific.[366] In this model it is impossible to estimate λ and the single parameters of U_{ij} simultaneously because such a model would be unidentified (Signorino, 1999, p. 284ff). Here, I follow the standard assumption according to which $\lambda = 1$; alternatively, the resulting estimates can be interpreted as the joint effect of cognitive capacity and utility parameter (i.e., $\lambda \times \beta$) (Signorino, 1999, p. 284). In general, this implies that a higher cognitive capacity or a lower level of randomness decreases the effect of all components in U_{ij}.

SOPHISTICATED VOTING

A sophisticated voter is not solely focused on their own payoffs but takes into account other subjects' choices and the logic of the collective decision problem. Thus, for sophisticated voting types their choice function is complemented with the core solution concept to ensure the theoretical consistency of actions and beliefs. The individual choice probability of alternative j then coincides with the probability of j being the core. Eqn. 6.8 presents the probability that actor i prefers alternative k over j in a pairwise contest.

$$p_{ikj} = \frac{exp\left(\lambda U_{ik}\right)}{exp\left(\lambda U_{ij}\right) + exp\left(\lambda U_{ik}\right)} \tag{6.8}$$

364 Goeree et al. (2002, p. 262) argued that the "logit probabilistic choice function" resembles a "convenient specification for empirical work".

365 In this simple form the denominator ensures that the probabilities of all nine alternatives add up to 1.

366 It follows that the estimations obtained display just a "representative subject" (Palfrey and Prisbrey, 1997, p. 835).

Yet, the core is defined at the collective level. Thus, the decisive question is with which probability alternative k wins a pairwise contest with alternative j when voted on in the group. For this, I calculate the probability by which each of the twenty-two theoretically feasible winning coalitions c would chose k over j. Eqn. 6.9 gives an example of this probability of a minimal winning coalition of the players 1, 2, 3 and 4 ($c = \{1,2,3,4\}$).

$$p_{1,2,3,4kj} = p_{1kj} * p_{2kj} * p_{3kj} * p_{4kj} \tag{6.9}$$

From Eqn. 6.9 it follows that the probability of alternative j to survive a pairwise contest with alternative k over all twenty-two possible winning coalitions c is equal to:[367]

$$p_{jk} = \prod_{c=1}^{22} (1 - p_{ckj}) \tag{6.10}$$

In other words, as long as one of the winning coalitions has a high probability of choosing k over j, the latter cannot be the core alternative because the core alternative beats every other alternative (i.e., it does not lose to any other alternative). Overall, the subjects could choose between nine alternatives $k = 1, 2, \ldots, 9$. Thus, in Eqn. 6.11 I multiply p_{jk} over all eight other alternatives $k \neq j$ in order to retrieve the probability with which alternative j wins every possible pairwise contest.

$$p_j = \prod_{k \neq j}^{9} p_{jk} \tag{6.11}$$

ADDING-UP CONSTRAINT

So far, Eqn. 6.11 states the predicted probability with which alternative j is adopted by a majority of subjects in the first round. The same holds true for sincere voting where the predicted probability results from the standard conditional logistic regression at the group level (cf. Eqn. 6.7). It is straightforward that the likelihood for reaching an agreement in this first round decreases with the size of the random disturbance.[368] Yet, the ballots of each round were iterated until a collective agreement had been reached. To accommodate this characteristic, I implement an adding-up constraint according to which the sum of the probabilities p_j over all nine alternatives equals one. This is shown in Eqn. 6.12.

$$\sum_{j=1}^{9} p_j = 1 \tag{6.12}$$

[367]The majority threshold of four out of six players has 15 minimal winning coalitions but is also reached by six possible coalitions of five and one unanimous coalition of all players.

[368]For $\lambda = 0$ each player acts randomly and throws an octahedral dice.

Sec. A.15 illustrates the predictions for the collectively chosen alternatives for different levels of random error for all payoff tables. The graphs are separated according to sophisticated and sincere voting. The resulting deviations are clearly visible; these differences were to be expected, as the two types are based on diverging models. However, the next section explains how they can still be integrated into a single model.

6.3.3 Mixture stage

My baseline assumption was that the subjects follow the theoretical solution concept of the core, i.e., they choose accordingly for sophisticated types. Yet, they are restricted by human cognitive capabilities. I expected at least some degree of sophisticated voting; in other words, I did not argue that all subjects voted in a sophisticated way all the time. Therefore, my statistical model composes a mixture stage which enabled me to estimate individuals' motivation while allowing for both sincere and sophisticated voting types.[369] In the end, the statistical model had to compute the probability with which each group collectively chooses any of the nine alternatives.[370]

If we take all of this together and use a subject's probability of sophisticated voting v_{il} similar to a weighting factor, a subject i's probability to vote for alternative j in a ballot l is then given by Eqn. 6.13. Here, $p^{sophisticated}$ refers to Eqn. 6.11 and $p^{sincere}$ to Eqn. 6.7.

$$p_{ijl} = v_{il} * p_{ijl}^{sophisticated} + (1 - v_{il}) * p_{ijl}^{sincere} \tag{6.13}$$

This leads to the question of what determines this probability of voting sophisticatedly. Sec. 6.2 argued that being part of the winning coalition might be a clue. Eqn. 6.14 defines the possible dependency of the degree of sophistication in a more general way. It shows a subject i's probability of voting sophisticatedly v_{il} in a ballot l and indicates possible co-variates h which may explain a subjects' degree of sophistication (cf. Tab. 6.1).

$$v_{il} = \frac{exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)}{1 + exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)} \tag{6.14}$$

CO-VARIATES

The range of possible variables which influence the extent of sophistication was restricted by the limited diversity of my sample. Of all subjects 92% were students and 88% of them were 25 years old or younger. Thus, it made no sense to use a "student-dummy" as well as "salary" or "age" as explanatory variables. Instead, I used the prior laboratory

[369] When discussing various theories of decision-making under uncertainty El-Gamal and Grether (1995, p. 1137) concluded that "current literature does not support the conclusion that subjects are sufficiently homogeneous to be described by a single theory."

[370] Cf. Redner and Walker (1984) for a methodological assessment of estimating mixture density parameters.

experience (i.e., the number of prior experiments the subject had participated in) as well as the previous knowledge of the subjects with respect to game theory. This knowledge was measured by four questions in the post-experiment survey: i) "Do you know the research area of game theory?", ii) "Do you know what prisoner's dilemma means?", iii) "Have you attended a lecture on experimental methods?" and iv) "Have you attended a lecture on micro-economics?"[371] The answers of the subjects were than normalized into an index reaching from 0 (i.e., four times no) to 1 (i.e., four time yes). Further, I include a winning coalition dummy (cf. Sec. 6.2). This is different from other co-variates which are determined exogenously. Yet, the coalition membership evolves endogenously in the course of the experiment.

Besides the sophistication, the amount of social considerations might also hinge upon exogenous factors. Therefore, I looked also into the explanatory power of co-variates at the utility level. More precisely, I interacted the parameters for inequality aversion (Eqn. 6.5) with potentially influential characteristics. Here, I could draw on a wide range of previous research. Ackert et al. (2009, p. 16) aptly summarized that "social behavior can be predicted by years of education, gender, university major, age, and primary household support. Participants who are less educated, female, non-business majors, older, and not primarily supported by their spouse are more likely to vote against their self-interest. Risk preferences are related to gender, but do not appear to be systematically related to age or altruistic behavior." As discussed earlier, the limited diversity of my sample restricts my choice of potential variables; of the mentioned aspects only gender and business major are feasible. Looking further into the experimental literature discloses an "unresolved debate about whether economists are different than other professionals" (Croson, 2005, p. 138).[372] Precisely because this has so far not been completely clarified, it makes sense not to be restricted to economics students but to enable volunteers from all university faculties to participate as subjects in experiments.[373]

The argument for the influence of gender on cooperation is supported by many experimental findings (for a comprehensive review cf. Croson and Gneezy, 2009). For example, Charness and Rustichini (2011, p. 77) found that "females cooperate significantly more than males". Another aspect of their work is even more interesting. The authors conducted prisoner's dilemma games and divided the participants into rooms. Then, subjects "played the game once with an audience of the same group ('at home') and once with an audience of the other group ('away')" (Charness and Rustichini, 2011, p. 77).

[371] As the experiment was conducted in Heidelberg these are translations of the original German questions.

[372] Croson (2005, p. 138) offered a literature overview of contributions which investigate the question of whether economists (or economics students) are different from other professionals (or non-economics students). On the one hand, they are found to free-ride more in social dilemmas (Marwell and Ames, 1981; Frank et al., 1993, 1996) and to offer less in ultimatum games (Carter and Irons, 1991). On the other hand, they are judged to be more cooperative in a lost-letter experiment (Yezer et al., 1996) and less likely to cheat compared to sociologists and political scientists (Laband and Biel, 1999). Also, Frey and Meier (2005, p. 170) argued that their measured "lower contribution of business economists [and not economists in general], compared to other students, is due to self-selection rather than indoctrination."

[373] Croson (2005, p. 138) assessed this "experimental economics practice [as] a sensible and conservative one."

I did not separate my participants into different rooms, but on a more general level this finding refers to the aspect of social distance which I discuss in Sec. 4.3.3. My subjects did not come from different pools and they did not play the game through various (technical or design-specific) means. Yet, an aspect which differed is how many other participants they knew (either just by sight or quite well). Here, more familiar faces will bring about a feeling of being at home. I operationalized this aspect by controlling the number of students within the laboratory at the same time who share the same major.

Overall, important aspects are the gender of the subject, whether or not the subject is an economics student and the number of fellow students from the same faculty in the laboratory. For the strategic component I used the number of prior participation in laboratory experiments and subjects' previous knowledge of game theory. In addition, I differentiate for subjects belonging to the winning coalition. Tab. 6.1 provides a descriptive overview of these co-variates; all associated computations are discussed in Sec. 6.4.2.

Table 6.1: Co-variates for the random utility mixture model

EXPLANATORY NOTE
The table depicts descriptive information for the co-variates used in the random utility mixture model estimation. The variables are separated according to utility and strategic level. For every co-variate I depict its average, SD, mean, minimal and maximal value. For gender a 0 indicates a male and a 1 a female participant. For an economics student and a winning coalition a 0 indicates false and a 1 indicates true. The variables fellow students and prior experiments show the amount of the respective variable. Theory knowledge displays its quantity normalized to the range of 0 (min) and 1 (max).

co-variates	N	mean	SD	median	min	max
UTILITY COMPONENT						
gender	168	0.57	0.50	1	0	1
economics student	155	0.46	0.48	0	0	1
fellow students	155	3.75	3.12	2	0	11
STRATEGIC COMPONENT						
winning coalition	1326	0.74	0.46	1	0	1
prior experiments	168	4.67	3.46	4	0	20
theory knowledge	168	0.48	0.33	0.5	0	1

STATISTICAL MODEL

Finally, by insertion of all the single elements in Eqn. 6.13 the derivation results in the likelihood function Eqn. 6.15 which is used for estimating the models. I apply standard maximum-likelihood techniques (following Goeree et al., 2002, p. 262) and use GAUSS' (Version 9.0.2, build 1114) constrained optimization algorithm (cf. Sec. A.1).

$$
\begin{aligned}
L\left(p_{ijl} \mid y_{ij}, x_{hil}\right) = {} & \frac{exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)}{1 + exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)} \\
& * \prod_{k \neq j}^{9} \prod_{c=1}^{22} \left(1 - \prod_{i \epsilon c} \frac{exp\left(\lambda U_{ik}\right)}{exp\left(\lambda U_{ij}\right) + exp\left(\lambda U_{ik}\right)}\right) \\
& + \left(1 - \frac{exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)}{1 + exp\left(\alpha + \sum_{h=1}^{H} \beta_h x_{hil} + \varepsilon_{il}\right)}\right) * \frac{exp\left(\lambda U_{ij}\right)}{\sum_{k=1}^{9} exp\left(\lambda U_{ik}\right)}
\end{aligned}
\tag{6.15}
$$

While the first two terms in the formula refer to sophisticated voting the following two terms express sincere voting. In both cases, the first expression indicates the relative weight of the respective behavior and the second expression represents the according probabilistic choice function. For sophisticated voting the probability depends on the prospect of an alternative j to be superior to any other alternative k over all possible winning coalitions c. For sincere voting I use standard conditional logistic regression.

6.3.4 Model specifications

This section briefly summarizes my theoretical model and explains four concrete specifications. The baseline model only comprehends purely self-interested actors (MODEL I). Starting from here, I supplemented my model with the three most relevant patterns of individual behavior which I discussed previously: sophistication, inequality aversion and risk aversion. All specifications are estimated using conditional logistic regression.

It is in particular difficult to filter out the sophisticated component of behavior. I accounted for this by including both sincere and sophisticated voting types (MODEL II). This followed the approach of Goeree et al. (2002) and resulted in a random utility mixture model (RUMM). The voting type distinction required a mixture stage that allows evaluating the extent to which the theoretical solution concept correctly predicts subjects' behavior. In other words, such a model contains two levels; one at which the utility of the subject was calculated (absolute payoff and relative payoff) and one which differentiates for the degree of sophistication with which this utility was maximized (e.g., choose highest payoff vs. choose highest payoff after weighting by its cue validity).

In addition, I incorporated further aspects into the strategic component of the model. Most importantly, I differentiated between the participants according to their affiliation to the winning coalition (MODEL III) as explained in Sec. 6.2.

The last model incorporated social preferences (MODEL IV) by implementing the inequality aversion concept of Fehr and Schmidt (1999). The big advantage of this approach is its ability to distinguish between advantageous and disadvantageous inequality.

Finally, one significant pattern of individual behavior is still missing. I was also concerned with the impact of subjects' risk attitudes as the design provided four different decision situations with different levels of uncertainty. This follows the call of Bradler (2009, p. 20) who emphasized that "research could make a further approach to merge theories of social preferences and theories on choice under risk together." Risk aversion extends to an individual's complete utility function. Thus, each model was estimated without (variant a) and with (variant b) risk aversion. The following section discusses the estimates of these eight models. In addition, I also consider co-variates and their explanatory power.

6.4 Results

While Chap. 5 discusses the experimental results by looking at the group performances, the following analysis investigates how those originated from individual choices. Of course, the insights found on the collective level (e.g., an increase of decision-making efficiency for delegation settings) remain valid. Yet, the individual data supplements the previous results with information on motive and sophistication of the votes.

As dependent variable, I used those individual choices which caused the collective outcome, i.e. the final ballots of every round. As I report parameter estimates for preferences based on last choices only, this follows the assumption that an individual's objectives remain "stable in the short run, [i.e.] for the duration of the experiment" (Bolton and Ockenfels, 2000, p. 171). This does not rule out that the "weights individuals give these objectives may well change over the long term, with changes in age, education, political or religious beliefs, and other characteristics" (Bolton and Ockenfels, 2000, p. 171).[374]

The computations were carried out for four different models: one which includes only self-interest (MODEL I), one that allows for a mixture of sincere and sophisticated voting (MODEL II), one that differentiates sophistication for members of the winning coalition (MODEL III) and one that applies the inequality aversion concept (MODEL IV). Each model was estimated without (variant a) and with (variant b) incorporating risk aversion.

The various models enable me, through their juxtaposition, to draw inference on subjects' motivation for their observed votes. In other words, I test the functional specification and perform (relative) validation tests. Thus, I am able to investigate the accuracy of the part-sincere-part-sophisticated hypothesis, to determine the relative importance of self-interest as compared to social preferences, whether risk aversion played a role, etc. I base the comparison on the significance of the parameter estimates, the likelihood of the statistical models and the Akaike information criterion (AIC, Akaike, 1973).[375]

6.4.1 Random utility mixture model estimates

This section discusses the conditional logistic regression model estimates for the various procedures. As explained in Sec. 4.2.6, the voting procedures exposed the participants to

[374]For a macro-analysis of data sets from various ultimatum game laboratory experiments cf. Prasnikar (1997). Most important, she found that people's perception of fairness is stable not only within one round but even with repeated play. Also, Borghans et al. (2008) examined the predictive power of personality and the stability of personality traits over the life cycle. They found that cognitive and personality traits evolve, but to different degrees and at different stages. As the important temporal dimension refers to the life cycle, it is at least clear that these traits are stable across situations.

[375]The AIC is a popular measure of the relative accuracy of fit of statistical models. It balances the degree of a model's variance and the extent of contained biases against each other (Burnham and Anderson, 2004). Based on the maximized likelihood function of a model, it assigns a penalty term for the number of included parameters; the "model with minimum AIC value is chosen as the best model to fit the data" (Bozdogan, 2000, p. 63). For a comprehensive discussion of its principles, asymptotic properties and inferential capabilities cf. Bozdogan (1987, 2000).

decision situations with different levels of uncertainty and complexity. Thus, looking at the (relative) importance of behavioral patterns within and across procedures, I am able to assign them to aspects of the decision-making process. For all procedures I list eight model specifications of which each comprises two sections. The upper section displays the parameter estimates of the utility component; i.e., self-interest, inequality aversion and risk aversion. The lower section displays the estimates for an actor's level of sophistication; i.e., the proportion of sophisticated voting subjects. I differentiate prominently for the proportion when accounting for the members of the winning coalition. Altogether, this yields complicated and diversified tables. To facilitate the contemplation, the tables display the estimated parameters and corresponding z-scores; this enables an importance comparison between different variables.[376] The score is calculated for every parameter as $z = \frac{estimated\,value}{standard\,error}$.[377] Thus, it measures by how many standard errors the estimated value deviates from 0. I used robust standard errors for all computations.[378]

It is crucial to note that the indicated degree of sophistication represents a special index. The given estimates are to be interpreted as values of the cumulative distribution function of the standard normal distribution. Specifically, an estimate ≤ -3 indicates 100% percent sincere voting while an estimate of ≥ 3 points to 100% sophisticated behavior. The values in between represent mixed shares; hence, at an estimate of 0 there is 50% of either type.[379] This representation has a crucial limitation. When displaying mixed voting behavior, i.e., some sophisticated as well as some sincere voting, it is not clear how much of the explanatory power is arising from either of the sophisticated or the sincere model. In other words, the data cannot tell which model explains the observations. Nevertheless, such an estimate still rejects the hypothesis that the mixture probability is 0. Thus, it proves the argumentation of Sec. 6.2 that the ballot is, in fact, composed partly of both sophisticated and sincere voting behavior.

As a baseline, I use the results obtained under POL which are shown in Tab. 6.2. MODEL I focused exclusively on self-interest which, unsurprisingly, turned out to be positive and significant. Including risk aversion shows that subjects were highly risk averse; i.e., small gains were relatively more important. With MODEL II the analysis moved beyond the standard conditional logit by implementing the mixture stage. This proves worth the effort as the model's explanatory power increased about as much when considering risk aversion. The estimates provided significant evidence of mainly sophisticated voting. Including sophistication and risk aversion together weakened the evidence and

[376] Assuming a standard normal distribution, the following z-scores and significance levels (two-tailed) are associated: $z = \pm 1.645 \longleftrightarrow p = 0.1$; $z = \pm 1.96 \longleftrightarrow p = 0.05$; $z = \pm 2.58 \longleftrightarrow p = 0.01$; etc.

[377] The actual calculation rule for z-scores is $z = \frac{observation\,value - sample\,mean}{standard\,error}$. Yet, as the logit model mean equals 0 I could apply the simplified version.

[378] With respect to robust standard errors, cf. Greene (2003, p. 267) for a general introduction, Croux et al. (2003) for a comprehensive discussion of examples, applications as well as scope of the approach and King and Roberts (2012) for an assessment of risks and problems of this method.

[379] Corresponding to the shape of the normal distribution this is no linear relationship. For example, a dispensation of 75% sincere and 25% sophisticated voting is reached at -0.675 while a distribution of 33.3% sincere and 66.6% sophisticated voting is reached at 0.415.

Table 6.2: Model estimates under pooling

EXPLANATORY NOTE
The table depicts the estimates of all eight logit models under the pooling procedure. As dependent variable, I used all individual choices of the final ballot of every corresponding experimental round. For all independent variables the table shows the beta coefficient in addition to the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

N = 456 individual votes	MODEL I		MODEL II		MODEL III		MODEL IV	
Variant of risk aversion	self-interest	self-interest and risk aversion	mixture	mixture and risk aversion	winning coalition	winning coalition and risk aversion	inequality aversion	inequality and risk aversion
UTILITY COMPONENT								
self-interest	0.095	5.331	0.165	1.775	0.160	0.376	0.075	0.199
	(13.069)	(3.822)	(33.850)	(2.175)	(18.171)	(41.330)	(6.972)	(20.557)
risk aversion		1.219		0.782		0.290		0.354
		(11.723)		(5.154)		(3.150)		(4.687)
D^+							0.046	0.088
							(5.848)	(49.111)
D^-							-0.067	-0.090
							(-8.753)	(-47.105)
STRATEGIC COMPONENT								
constant			2.834	0.784	0.656	0.434	1.150	0.8631
			(38.786)	(1.994)	(3.787)	(24.260)	(2.322)	(40.613)
winning coalition					8.625	4.146	7.712	4.894
					(40.721)	(83.914)	(9.128)	(74.808)
Log likelihood	-958.18	-878.16	-877.32	-858.98	-844.44	-832.59	-813.11	-805.59
Pseudo R^2	0.04	0.12	0.12	0.14	0.16	0.17	0.19	0.20
AIC	1918.36	1760.32	1758.64	1732.96	1694.88	1673.18	1636.22	1623.18

effect of both. Yet, they stayed significant and retained their estimates. Next, MODEL III investigated the sophistication of members within and outside the winning coalition separately. Here, the estimated probability of encountering sophisticated voting types was around 74% for subjects outside and almost 100% for subjects within the winning coalition. Also, the level of revealed risk aversion dropped significantly compared to the previous models. Finally, MODEL IV looked into the parameters for inequality aversion. The subjects disliked getting less than others (negative D^-), while at the same time they revealed a highly significant preference for getting more than others (positive D^+). In other words, the subjects exhibited a highly competitive behavior as they were motivated by relative gains; this stays significant when controlling for sophistication as well as for risk aversion. Across all model specifications, the explanatory power clearly improved when adding risk aversion. The increase was highest for the simplest model; for more multifaceted specifications the value of adding further parameters was less extensive. Yet, under POL the best performing specification was MODEL IV which includes the mixture component as well as inequality and risk aversion.

That subjects were motivated by relative gains contradicts some of the previous work on social preferences where altruistic behavior in collective decisions was identified (e.g., Goerg et al., 2007; Levin, 2009). Yet, those findings were mostly obtained in other experimental designs such as the ultimatum game, dictator game, trust game, gift exchange game or public good game (cf. Sec. A.14). These are games in which a subject's competitiveness (i.e., the preference for advantageous inequality) does not alter their equilibrium behavior (Fehr and Schmidt, 1999, p. 850). My design placed the participants into a different context. Goeree et al. (2002, p. 267) pointed out that the decision context affects an individual's preferences as in bargaining situations "the focus is more clearly on issues of division", which makes altruism less important and inequity aversion more salient. In line with my results, Dittrich and Ziegelmayer (2010, p. 9) also found a lack of aversion to advantageous inequality in their bilateral gift exchange game.

Next, I looked at the model estimates for the second stage under SEQ. This resembled a simpler version of the decision situation under POL as the number of players and alternatives declines. The results are shown in Tab. 6.3. When considering only self-interest (MODEL I) they appeared rather similar to POL. One notable difference was less risk aversion and that the increase brought by it to the explanatory power was rather small. This changed once the mixture stage of MODEL II allowed for sincere and sophisticated voting types. The estimated probability of encountering sophisticated voting was 25%, and 13% when controlling for risk aversion. The dominance of sincere voting also prevailed in both groups ($\geq$90%) when accounting for insiders and outsiders of the winning coalition separately in MODEL III. In general, including the strategic component improved the model significantly compared to the simplest version. MODEL IV completed the analysis by looking at inequality aversion but the knowledge gained was very little. The only significant parameter indicated a rejection of advantageous inequality. Yet, this disappeared

Table 6.3: Model estimates of the second stage under sequential delegation

EXPLANATORY NOTE
The table depicts the estimates of all eight logit models for the second stage under the sequential delegation procedure. As dependent variable, I used all individual choices of the final ballot of every corresponding experimental round. For all independent variables the table shows the beta coefficient in addition with the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

N = 180 individual votes	MODEL I		MODEL II		MODEL III		MODEL IV	
Variant of risk aversion	self-interest	self-interest and risk aversion	mixture	mixture and risk aversion	winning coalition	winning coalition and risk aversion	inequality aversion	inequality and risk aversion
UTILITY COMPONENT								
self-interest	0.080	0.347	0.467	1.141	1.287	1.757	0.842	1.605
	(7.234)	(2.192)	(5.333)	(0.623)	(2.244)	(1.738)	(2.031)	(1.350)
risk aversion		0.563		2.799		1.191		0.902
		(3.697)		(4.925)		(1.657)		(4.217)
D^+							-0.743	0.065
							(-2.258)	(0.195)
D^-							-0.080	-0.497
							(-0.248)	(-0.456)
STRATEGIC COMPONENT								
constant			-0.618	-1.120	-1.444	-1.609	-1.525	-1.431
			(-2.443)	(-3.528)	(-3.056)	(-2.947)	(-3.112)	(-2.991)
winning coalition					-0.373	-1.091	-1.313	-0.541
					(-0.683)	(-1.609)	(-2.559)	(-0.985)
Log likelihood	-161.85	-159.36	-145.77	-135.90	-133.01	-125.21	-126.99	-124.32
Pseudo R^2	0.18	0.19	0.26	0.31	0.33	0.37	0.36	0.37
AIC	325.70	322.72	295.54	277.80	272.02	258.42	263.98	260.64

Table 6.4: Model estimates of the first stage under sequential delegation

EXPLANATORY NOTE
The table depicts the estimates of six logit models for the first stage under the sequential delegation procedure. As dependent variable, I used all individual choices of the final ballot of the first stage of every corresponding experimental round. As only for this stage is a division by beliefs necessary, the respective two variants are shown in Tab. 6.5 separately. For all independent variables the table below shows the beta coefficient in addition to the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

N = 180 individual votes	MODEL I		MODEL II		MODEL III	
Variant of risk aversion	self-interest	self-interest and risk aversion	mixture	mixture and risk aversion	winning coalition	winning coalition and risk aversion
UTILITY COMPONENT						
self-interest	0.057	1.150	0.104	0.552	0.110	0.857
	(6.940)	(1.600)	(6.055)	(7.229)	(5.820)	(2.746)
risk aversion		0.886		0.530		0.670
		(3.784)		(3.282)		(2.021)
D^+						
D^-						
STRATEGIC COMPONENT						
constant			-0.448	-0.584	-2.358	-4.124
			(-1.322)	(-1.675)	(-2.347)	(-4.857)
winning coalition					0.452	0.491
					(0.418)	(0.506)
Log likelihood	-164.26	-155.65	-155.71	151.86	-149.01	-142.51
Pseudo R^2	0.17	0.21	0.21	0.23	0.25	0.28
AIC	330.52	315.30	315.42	309.72	304.02	293.98

when controlling for risk aversion. It seems that subjects abstained from choosing very profitable alternatives because they feared that those allow no consensus. Across all specifications for the second stage under SEQ, the best performance was achieved by MODEL III, which includes the mixture component and risk aversion.

Tab. 6.4 stays with SEQ but turns to the decision-making of the first stage under this procedure. Here, players have to consider the two other players of their own group as well as the subsequent choices of the three players in the second group. Thus, they were exposed to a higher level of uncertainty when making their choices. MODEL I showed that self-interest and risk aversion played an important role. This also holds true for all other model specifications under this procedure. The estimates of the strategic component in MODEL II and MODEL III resembled the results of the second stage. Sincere voting was predominant ($\geq$67%) and differentiating for the winning coalition did not change this but rather made the lack of sophistication even more explicit ($\leq$8%). Looking at the two stages of SEQ together, it becomes obvious that risk aversion and sincere voting dominated both. Here, the CRRA parameter as well as the degree of sophistication (across all subjects) was stronger in the second stage. This high risk aversion is surprising as I characterized this decision situation with a low level of uncertainty (cf. Tab. 4.2)

So far, MODEL IV is missing. In Sec. 6.3.2 I stated that the utility parameters and λ were not estimated to be actor-specific as such a model would be unidentified (Signorino, 1999, p. 284ff). Thus, the results in Tab. 6.4 are based on the assumption that subjects in both stages possess the same parameter estimates. In other words, subjects in the first stage believed that their colleagues in the second stage would share their motivation. This might be considered as a rather strong and unrealistic assumption (henceforth this assumption is referred to as *strong beliefs*). Tab. 6.5 contrasts these beliefs with a second variant that assumed that players at the first stage believe that all remaining alternatives are equally likely to be elected in the second stage. To implement this, I set all second stage parameters equal to zero (henceforth this assumption is referred to as *flat beliefs*).[380]

As in the second stage, MODEL IV without risk aversion indicated that subjects dislike advantageous inequality. This effect became again insignificant when accounting for risk aversion. The differences resulting from the two belief assumptions are not so trivial. With flat beliefs, subjects showed distaste for relative losses and risk aversion was less important. Also, the differentiation for belonging to the winning coalition turns out to be significant. Yet, this modified the supremacy of sincere voting only in the second decimal. Across all models for the first stage under SEQ the best performing model was MODEL IV based on flat beliefs and including the mixture component, risk aversion and inequality aversion. The better performance of this belief assumption corresponds to the obtained parameters. As most voting was sincere, the assumption of flat beliefs corresponds best to the underlying focusing of subjects on themselves.

[380] I restrict the comparison of different beliefs to the fully specified model in order not to overload the analysis. However, the same argument could be made for each specification.

Table 6.5: Model estimates under sequential delegation separated for beliefs

EXPLANATORY NOTE
The table depicts the MODEL IV estimates for the first stage under the sequential delegation procedure for different beliefs. As dependent variable, I used all individual choices of the final ballot of the first stage of every corresponding experimental round. For all independent variables the table below shows the beta coefficient in addition with the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

N = 180 individual votes	MODEL IV		
Variant of risk aversion	inequality aversion	inequality and risk aversion	
Beliefs	strong	strong	flat
UTILITY COMPONENT			
self-interest	0.2911	0.667	0.2918
	(3.127)	(3.038)	(3.871)
risk aversion		0.570	0.288
		(1.939)	(4.181)
D^+	-0.1827	-0.031	0.097
	(-2.327)	(-0.318)	(1.279)
D^-	0.056	-0.052	-0.280
	(0.769)	(-0.553)	(-3.604)
STRATEGIC COMPONENT			
constant	-4.581	-4.584	-4.334
	(-5.421)	(-5.421)	(-5.919)
winning coalition	0.234	0.234	1.023
	(0.370)	(0.370)	(5.304)
Log likelihood	-141.81	-141.81	-138.05
Pseudo R^2	0.28	0.28	0.29
AIC	295.02	295.62	288.10

Finally, Tab. 6.6 lists the estimates for SIM. Due to the high amount of uncertainty and complexity under this procedure I obtained highly significant estimates of risk aversion in all four specifications. MODEL II and MODEL III indicated the from SEQ familiar pattern that sincere voting prevailed and that the affiliation to the winning mattered only little. Looking at MODEL IV, I found no evidence for social preferences at all. The knowledge gained when moving from the simplest to the most diverse model is lowest under SIM in comparison to the other procedures. Almost the full extend was already reached by adding risk aversion. However, on consideration of likelihood and AIC, the best performing specification was MODEL III which includes the mixture component and risk aversion.

6.4.2 The influence of co-variates

Sec. 6.3.3 discusses possible influencing factors which might explain why a subject acted socially and sophisticatedly or not. Yet, the concrete expectations for my specific design are not clear. On the one hand, substantial literature discusses influential properties found in many studies. On the other hand, on examination of the most similarly constructed experiments, e.g., Goeree et al. (2002) and S&K (2010), no influence of co-variates was observed.[381] Moreover, additional variables further increase the intricacy of the objective function to be estimated. With respect to my number of observations, this

[381] More precisely, Goeree et al. (2002) found on average that altruism parameters were the same for men and women. However, they observed that male altruism was significantly more dispersed.

Table 6.6: Model estimates under simultaneous delegation

EXPLANATORY NOTE
The table depicts the estimates of all eight logit models under the simultaneous delegation procedure. As dependent variable, I used all individual choices of the final ballot of every corresponding experimental round. For all independent variables the table shows the beta coefficient in addition to the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

N = 510 individual votes	MODEL I		MODEL II		MODEL III		MODEL IV	
Variant of risk aversion	self-interest	self-interest and risk aversion	mixture	mixture and risk aversion	winning coalition	winning coalition and risk aversion	inequality aversion	inequality and risk aversion
UTILITY COMPONENT								
self-interest	0.062	0.998	0.134	1.025	0.127	1.087	0.113	1.193
	(12.216)	(3.932)	(6.599)	(3.510)	(6.956)	(50.785)	(2.194)	(24.297)
risk aversion		0.864)		0.789		0.806		0.982
		(9.757)		(7.710)		(61.550)		(31.458)
D^+							-0.053	-0.003
							(-1.182)	(-0.256)
D^-							-0.052	0.000
							(-1.239)	(0.027)
STRATEGIC COMPONENT								
constant			-0.511	-1.811	-1.382	-4.091	-6.602	-4.389
			(2.186)	(3.531)	(-3.011)	(-10.386)	(-10.654)	(-88.667)
winning coalition					-0.171	-0.966	-0.990	-1.122
					(-0.337)	(-1.407)	(-1.462)	(-22.669)
Log likelihood	-448.86	-416.68	-431.58	-412.02	-427.49	-407.37	-410.34	-407.12
Pseudo R^2	0.20	0.26	0.23	0.26	0.24	0.27	0.27	0.27
AIC	899.72	837.36	867.16	830.04	860.98	822.74	830.68	826.24

should not be overdone. However, in this section I look into the influence of co-variates for two reasons. Firstly, because of the possibility of inductive findings and, secondly, because this step resembles a robustness test for the results obtained in my experiment; and therefore it increases their validity.

In accordance with the previously discussed literature the co-variates were added into both parts of the statistical model. For the strategic component these were the amount of prior participation in laboratory experiments (prior experiments) and subjects' knowledge of game theory and experimental methods (theory knowledge). The utility component accounts for the gender of the subject and either if the subject is an economics student (econ student) or the number of fellow students of the same faculty in the laboratory. Here, the co-variates were included in the form of interaction effects with the inequality aversion coefficients. Starting from the baseline of a male, non-economics student (or a subject without fellow students) with no prior participation in experiments and no theory knowledge, the co-variates indicate how a change in one of these aspects affected social behavior or strategic voting; i.e., they enable the comparison between subgroups of participants. The results are shown in Tab. 6.7.[382]

Under POL I obtained the same findings as before; self-interest was highly relevant, subjects behaved competitively (they liked advantageous and loathed disadvantageous inequality) and sophisticated voting was significant. Neither of the co-variates influenced the results. The differences between specifications for either economics students or fellow students seem negligible. In fact, the statistical measures of the models' predictive power rarely change between any procedure for these two co-variates. Looking at the utility component of the delegation procedures, I obtained significant estimates for subjects' characteristics. In the second stage SEQ I found already familiar patterns (subjects disliked relative gains) but also a significant influence of the number of fellow students. Interestingly, a larger number of colleagues with the same field of study amplified the rejection of disadvantageous inequality. Yet, the results in the second stage of SEQ were obtained from a reduced data set. When using all observations, the logit model was "completely determined" due to "hidden collinearity" within my co-variates (cf. Sribney, 1999).[383] Here, a specific pattern across the co-variates had only one outcome. There are various strategies to deal with this problem as, e.g., omitting the variable responsible, centering the predictors or excluding the affected observations (cf. Kleinbaum et al.,

[382] I restrict the investigation of co-variates to subjects' social behavior and strategic voting in order to maintain a comprehensible analytical framework. Manifold experimental studies have looked into the influence of personal aspects on risk aversion (cf. Cox and Harrison, 2008; Eriksson and Simpson, 2010). While I included this behavioral pattern in my utility model (cf. Sec. 6.1.5), it is not my main research objective to identify the impact of personal characteristics on risk aversion. Current contributions also discuss possible systematic relationships across distinct aspects of individual preferences such as risk aversion, social preferences, reciprocity, etc. (e.g., Ackert et al., 2009). However, these considerations are beyond the scope of my study.

[383] The symptoms are similar to those of multicollinearity as, e.g., the co-variates have insignificant coefficients when entered jointly in the regression, but each has a significant coefficient when entered individually.

Table 6.7: Model estimates including co-variates

EXPLANATORY NOTE

The table depicts the influence of co-variates on social preferences and sophisticated behavior in all four decision situations. As dependent variable, I used all individual choices of the final ballot of every experimental round. The included co-variates for the strategic component are the number of prior participation in laboratory experiments (prior experiments) and subjects knowledge of game theory and experimental methods (theory knowledge). For the utility component two variants are considered; once the gender of the subject and if the subject is an economics student (econ student) and once gender together with the number of fellow students in the laboratory (fellow students). For all independent variables the table shows the beta coefficient in addition to the corresponding z-scores based on robust standard errors in parentheses. The pseudo R^2 is calculated following McFadden (1974).

	POL		SEQ second stage		SEQ first stage		SIM	
	N = 456		N = 168		N = 180		N = 480	
UTILITY COMPONENT								
self-interest	0.069	0.065	1.326	1.278	0.181	0.242	0.121	0.120
	(4.935)	(5.156)	(3.334)	(3.194)	(1.437)	(2.115)	(2.318)	(2.298)
D^+	0.046	0.045	-1.079	-1.036	-0.187	-0.264	-0.080	-0.076
	(3.564)	(3.693)	(-3.391)	(-3.187)	(-1.663)	(-2.411)	(-1.723)	(-1.619)
$D^+ \times$gender	-0.001	-0.001	0.121	0.116	0.159	0.124	0.040	0.039
	(-0.111)	(-0.065)	(1.534)	(1.397)	(2.551)	(2.764)	(2.267)	(2.133)
$D^+ \times$econ student	0.010		-0.097		0.126		-0.003	
	(0.840)		(-1.195)		(2.284)		(-0.154)	
$D^+ \times$fellow students		0.002		-0.075		0.025		-0.005
		(0.918)		(-1.015)		(2.424)		(-0.265)
D^-	-0.067	-0.075	0.528	0.632	-0.074	-0.064	-0.031	-0.032
	(-5.484)	(-6.135)	(1.453)	(1.917)	(-0.463)	(-0.455)	(-0.675)	(-0.629)
$D^- \times$gender	0.001	0.003	0.196	0.131	0.023	0.012	0.010	0.021
	(0.061)	(0.280)	(0.748)	(0.756)	(0.188)	(0.160)	(0.326)	(0.626)
$D^- \times$econ student	-0.015		-0.027		0.063		-0.066	
	(-1.118)		(-0.129)		(0.545)		(-1.865)	
$D^- \times$fellow students		0.001		-0.069		0.024		-0.010
		(0.317)		(-1.864)		(1.181)		(-1.640)
STRATEGIC COMPONENT								
constant	4.446	4.590	-1.1543	-1.021	-0.137	-0.266	-0.908	-0.915
	(4.359)	(4.297)	(-1.189)	(-1.204)	(-0.222)	(-0.455)	(-1.203)	(-1.607)
prior experiments	-0.092	-0.123	-0.094	-0.182	-0137	0.015	-0.013	-0.002
	(-1.058)	(-1.376)	(-0.417)	(-1.369)	(-0.059)	(0.196)	(-0.113)	(-0.120)
theory knowledge	-0.668	-0.042	0.406	0.423	-0.342	-0.455	-1.740	-1.733
	(-0.602)	(0.378)	(0.284)	(0.394)	(-0.374)	(-0.487)	(-1.469)	(-1.573)
Log likelihood	840.27	840.03	121.20	119.96	140.06	140.48	408.55	408.79
Pseudo R^2	0.16	0.16	0.39	0.39	0.29	0.29	0.27	0.27
AIC	1704.54	1704.06	266.40	263.92	304.12	304.96	841.10	841.58

2007, p. 365ff). I chose the last option; after identifying the responsible pattern based on the predicted logits, I excluded all observations which comprised $prior\,experiments = 0$ and $theory\,knowledge = 0$. These were 12 observations and applied to four of the 168 participants.

In the first stage of SEQ I found that subjects disliked advantageous inequality. This behavior was stronger for men compared to women, for non-economics students compared to economics students and the less fellow students took part in the same experiment. These insights correspond partially to the results under SIM; here, subjects also rejected advantageous inequality and this effect was weaker for women. Furthermore, economics students refused disadvantageous inequality to a greater extent.

In summary, the distinction between the economics student and fellow student co-variate is not decisive as models' predictive accuracy rarely changed. When focusing on co-variates of the strategic component, I observed no differences for the degree of sophistication. This does not mean that a combined effect of the variables cannot be significant, but clearly the differences between subgroups were not. Looking at the co-variates of the utility component, I found some effects but no general pattern that holds true across all procedures.

6.5 Chapter summary

This chapter analyzes the experimental data on the individual level. I start in Sec. 6.1 with an introduction to common patterns of individual behavior found in prior studies. As many of them were discovered and validated in laboratory experiments, it was likely that they would also play a role in my design. Sec. 6.2 addresses the limitations of humans' cognitive capacities (Simon, 1957) and explains that an experiment such as mine has to account for both sincere and sophisticated voting subjects (Goeree et al., 2002). Next, Sec. 6.3 brings these insights together and derives the statistical model. It includes various specifications to cover the different behavioral patterns, contains a mixture stage to distinguish for sophistication and comprises error terms to filter out the noise of collective decisions. Overall, I specified eight model variants which are compared in Sec. 6.4 by juxtaposition in order to determine the impetus for an individual's behavior. Not decisive were various co-variates which I tested for their influence on social considerations as well as sophistication.

If one looks across the models it becomes clear that decision makers' substantial motivation must be modeled with a view to their risk aversion. A model's predictive accuracy increased when including CRRA; this pattern was constant across all variations of group size and decision rule. The importance of risk aversion in an environment characterized by uncertainty is no new finding. Yet more important, also in collective decisions risk aversion has to be modeled on the individual level.

A second element that enhanced model performance across all specifications was the consideration of sincere and sophisticated voting behavior at the same time. The implementation of the mixture stage (cf. Sec. 6.3.3) allows the estimation process to allocate the behavior observed to the corresponding concept. This holds true even when sincere voting is almost exclusively present as it facilitates the identification of noise within the collective agreement. Overall, sophisticated voting took place predominantly only under POL. This is most evident when compared to SIM, which was constituted entirely of sincere choices. Here, a high level of uncertainty in combination with a high level of complexity prevented subjects from evaluating the conceptualized solution concept. As the computations became too demanding, they stopped playing in a sophisticated manner. POL also resembles the only procedure under which a distinction for the extent of sophistication based on participation in the winning coalition turned out to be significant. Of the two stages under SEQ, the second disclosed more strategic voting even if only at a low level. Thus, the ability for sophisticated behavior decreases according to the level of uncertainty.

In my experiment, I found that subjects were motivated by competition rather than by fairness-related considerations. First of all, subjects cared about their own payoff and were risk averse with respect to it. The aversion was highest in the second stage of SEQ where subjects seem solely concerned about avoiding the lowest payoff. Secondly, subjects disliked relative losses. The disapproval is significant under POL and in both stages of SEQ. Yet, it was most pronounced in the first stage where subjects were primarily afraid of earning less than others. At the other extreme, under SIM neither relative gains nor losses are significant; i.e., subjects disregarded other players and stayed purely self-interested. Thirdly, under POL subjects' disclosed preferences for advantageous inequality. My results are therefore clearly not in line with the expectations of the Fehr and Schmidt (1999) inequality aversion model.

When the insights into strategic voting and social considerations were linked, the prevalence of sincere behavior and the distaste of inequality coincided. The best performing model in both stages of SEQ as well as under SIM was MODEL III, which does not include social preferences. Yet, their consideration very well enhanced the explanatory power of the models under POL. It seems that social preferences may be a luxury which subjects only take into account when they are reasonably certain about, e.g., other players' motivations, choices, and, accordingly, the resulting consequences of their own votes. Goeree et al. (2002, p. 267) showed that the decision context affects an individual's preferences. Yet, for this insight to be maintained it is foremost essential that the encountered task must be logically comprehended. Uncertainty causes subjects to focus more heavily on their immediate self-interest. Granted, looking only at one's own payoff is less demanding while considering others requires more complex thoughts. So, complexity might also drive the degree of self-focusing. Yet, a fair solution (e.g., a uniformly distributed alternative) has also focal point qualities which help players to coordinate under any circumstances.[384] In my experiment such an outcome (as well as unanimity, cf. Sec. 5.4) was rarely achieved even at low levels of complexity.

In particular, the high degree of risk aversion observed in the second stage of SEQ is puzzling. This task was the simplest one my subjects encountered. They were not left uncertain about further developments and, thus, the level of uncertainty was also low. Looking at the two characteristics which I used to distinguish the four decision situations (cf. Tab. 4.2), I must judge my rating system as an approach which is too static. It turns out that I did not take (sufficient) account of a specific source of uncertainty when developing my design, i.e., the unification process. The collective decision-making in the second stage of SEQ was simple and handled quickly; Tab. 5.14 revealed that this task hap the highest efficiency across all procedures. This left subjects without any information about the other players as most rounds ended with the first ballot. In contrast, under POL most collective decisions required at least four ballots. This enabled the subjects to reach a deeper understanding of the occurring problem.[385]

Overall, the findings suggest that uncertainty and the complexity of a task reduced subjects' ability to act in a sophisticated way and undermined the relevance of social preferences. Thus, nonseparability in conjunction with the respective institutional arrangement indeed influenced individual behavior. More precisely, subjects who faced a difficult setting disregarded others and focused on their own material payoff. This was aptly summarized by Cabrales et al. (2010, p. 2276), who stated that "where strategic uncertainty conflicts with social preferences in terms of their respective recommendations [...] the former seems to be subjects' primary concern." Every aspect that blurs the ability to reason diminishes social considerations. By contrast, subjects who faced less demanding challenges were more likely to compare their own to other players' payoffs. Thus, using the means of delegation must maintain players' ability to understand the impact of their individual choices on the final outcome. If not, they will focus on their own self-interest and disregard others because taking them into account would make an evaluation of the situation even more complicated.

Furthermore, a consideration of the payoffs of co-players does not automatically imply that subjects become socially engaged. In my experiment many behaved competitively. Yet, aiming to be more successful is also an action requiring one to consider others. I do not judge whether a competitive attitude is good or bad. Competition can serve as a discovery procedure and lead to welfare levels not achievable through collusion.[386] However, this conclusion should be independent from the direction of inequality or the investigated setting. In fact, social preferences under majority rule seem unpredictable;

[384] Colman (2003, p. 139) pointed out that "focal point selection in pure coordination games is inexplicable, though it is easily achieved in practice".

[385] As discussed in Sec. 4.3.1, the course of the unification process is an exciting puzzle in itself. Further research (cf. Sec. 9.2.3) might look thoroughly into this matter, which is beyond the limits of this study.

[386] This is aptly demonstrated by Kirstein and Schmidtchen (2003) in classroom experiments on the invisible hand argument of Adam Smith (1776).

at some point true fairness as exhibited in two-player games turns into competitiveness. Consequently, the findings recommend that real-world decision rules should be designed to reduce delegates' uncertainty about the political consequences of their choices. In order to preserve the relevance of social preferences, single decision makers should not only be experts in the subjects at hand, but they should also possess sufficient information on the preference distribution and the strategies of their co-deciders.

7 Post-experiment survey

I conclude my evaluation of the laboratory experiment with this chapter. The following sections present additional information which I collected in a mandatory post-experiment survey after the voting experiment. In the questionnaire I asked the participants for some personal information, such as profession, subject and semester of study (if they were students), age, gender, etc.[387] Overall, the main part of the survey focused on the decision-making of the participants during the experiment. The questionnaire asked subjects whether their decisions made in the experiment had been guided by a specific rule, questioned them about their focus in the process of reaching a decision, and about how they carried out the payoff comparisons, as well as about their criteria for evaluating competing alternatives. For all these issues the questionnaire provided 'yes or no' answers as well as free-input fields.

This short survey could not solve all unanswered questions, but it provided useful information, most importantly with respect to the validity of the experimental design. The participants were asked to describe and classify their decisions; this revealed the participants' assessment of whether and how nonseparability affected their behavior under each institutional arrangement. In combination with the voting data this delivered fruitful insights. Also, using an inductive approach to look into the questionnaire led to new aspects of the analysis.

The following sections contain the empirical analysis of the subject's responses. I start in Sec. 7.1 with some general comments the participants made about the experiment as a whole. Then Sec. 7.2 analyzes subjects' entries in the binary questionnaire. Next, I turn towards the free-input fields, in which the participants described "with their own words" their decision-making process. I analyze the data using a qualitative content analysis following Mayring (2002, 2010) in Sec. 7.3. Furthermore, I use in Sec. 7.4 the text scaling software WORDFISH developed by Proksch and Slapin (2009), to perform a computational text analysis. Finally, Sec. 7.5 summarizes the results of this chapter.

7.1 General comments

In laboratory experiments it is common practice to give participants the opportunity to comment on the experiment; either while answering a post-experiment survey or when

[387]The descriptive information is discussed in Sec. 4.4.

receiving their payoffs. Yet, often subjects might not want to give an elaborate commentary. This may depend on the duration of, interest in, or excitement about the experimental task. And sometimes subjects might provide a lot of remarks that have nothing to do with the experiment at all. Fortunately, I did not experience anything of this kind during my investigations.

In my experiment all subjects were given the voluntary option to comment on any aspect of the experiment by typing remarks in a free-input field.[388] Overall, 49% of the participants took advantage of this opportunity. The statements received are concerned with the overall duration of the experiment and the waiting times within the sessions. In addition, the participants asked many questions about communication opportunities.

I conducted an experiment on collective decision-making. Such a collective decision was only reached when enough individual players agreed on an alternative. As I reshuffled the groups after every round, the fastest group always had to wait until all remaining groups had agreed as well. The same principle holds true for many individual experiments because re-allocating takes time. But a collective agreement results in delays larger than in individual tasks. At the beginning of every experimental session this was explained to the participants in order to appeal to their patience. Nevertheless, the waiting time between the experimental rounds was the main object of criticism. About 42% of all comments (i.e., 21% of all participants) expressed such complaints. While the delays were an issue across all procedures, the strongest rejections are related to POL. As shown in Sec. 5.4, under this procedure it took the longest for participants to reach an agreement. For future experimental contributions this might be a valuable key insight to prevent participants from this sort of negative experience. Possibilities are, for example, to use smaller groups, providing distracting screens or tasks, disbanding group reshuffling, etc.

Admittedly, with sessions lasting 1:10 to 2:45h, the experiment was at the upper limit with respect to the overall duration. This was expressed in some of the comments which stated that the sessions "took quite a while". Overall, 15% of all comments were related to this topic; however, collective tasks need time, sometimes much more than individual ones. No experimental session exceeded the time duration announced in the recruiting email send to all participants.

During my whole experiment, communication between the subjects was forbidden. Yet, as subjects were engaged in a collective task they missed this opportunity. Interestingly, the participants did not complain about this in the questionnaire. The desire for communication was rather expressed verbally to the experimenter when collecting their payoffs.[389] Participants wanted to speak their mind to others about their (from their point of view not understandable) behavior, to coordinate collaboration, to scrutinize the interest in others' thoughts, etc. The reasons to ask for overall silence are still coherent

[388] More specifically, the question asked "Is there something left you want to tell us?" (Translation of the originally German question).

[389] I did not record the payment process, so I can only state a rough estimate: to the best of my knowledge approximately 30% of the participants asked about communication and why it was forbidden.

(cf. Sec. 4.2.2), but the clarity and frequency with which the participants desired conversation opportunities with others makes me recommend it for future experiments. Especially in the context of the classification between sincere and sophisticated behavior it is interesting how participants would make use of their ability to communicate. In Sec. 9.2.3 I discuss in detail the possibility of using communication in a much more pivotal role.

Finally, experimental results can be distorted by EDE (cf. Sec. 4.2.2), i.e., a change in subject behavior according to what they think is appropriate behavior (Zizzo, 2010, p. 2). The comments allowed to look for evidence of this. For example, this includes statements like that the underlying principle was discovered, that the participant was not deceived, that the expected solution technique had been applied, that the participant was able to uncover the intentions of the experimenter, etc. I found no such indications of EDE in the comments of the subjects.

7.2 Binary questionnaire

The binary questionnaire restricted subjects to answering with "yes" or "no". Of course, subjects could also opt to no respond (respectively not "clicking" on either of the corresponding check-boxes), but this did not occur at all in this part of the survey. Tab. 7.1 shows the questions and answers of the questionnaire.[390]

Table 7.1: Results of the binary questionnaire

N = 168	YES	NO
"Did you follow a specific decision rule?"	87%	13%
"Which players did you focus on?"		
myself	99%	1%
myself and others	88%	12%
others	1%	99%
"As the group was split up, on which players did you focus on?"		
myself	65%	35%
myself and the members of my delegation	46%	54%
the members of my delegation	4%	96%
all six players	38%	62%
"Which criterion of an alternative did you consider?"		
the amount of my payoff	73%	27%
an equal distribution of the payoffs	34%	66%
the sum of all payoffs	24%	76%
the difference between the lowest and highest payoff	22%	78%

The results reveal that 87% of the participants answered the most basic question, if they "followed a specific decision rule", with yes. After all, that 13% did not follow a rule is

[390] The questions in Tab. 7.1 are translations of the originally German questionnaire.

surprising, as the question does not exclude simplest rules. When I asked for a subject's focus in the process of payoff comparison, almost all subjects took their own payoffs into consideration (99%). Yet, at least 88% also paid attention to other players, and 1% focused solely on others. Thus, as expected, one's own payoff was most important. Under either delegation procedure the pattern becomes less explicit. When the group was split up, still a majority of 65% focused on themselves, but 46% took their delegation members into account, and 38% considered all six players. Only 4% focused solely on the three members of their delegation. These numbers suggest that altering the decision-making mechanism resulted in subjects shifting their focus to some extent from their own towards other people's payoffs. The last question asked for the criteria by which subjects evaluated the payoffs. A majority of 73% focused on the amount of their own payoff. Approximately a quarter of all participants paid attention to the sum of all payoffs (24%) or the difference between the lowest and highest payoff (22%). About 34% looked out for an equal distribution. While a subject's own success still dominated the individuals' attention, this verifies social considerations. Of course, considering others and finally also voting according to their benefit may be two different things. But to look at relative criteria is a first step towards "fairness".

Tab. 7.2 shows the amount of different criteria a subject considered. 55% of participants stated just one criterion. 33% percent considered two and 10% three criteria when making their decision. The extremes seem negligible. Only 1% considered either none or four different criteria.

Table 7.2: Number of different criteria a subject considered

N = 168					
# of criteria	0	1	2	3	4
% of participants	1%	55%	33%	10%	1%

Overall, 44% of the subjects considered more than one criterion at a time. Tab. 7.3 takes a look at the correlation between the different criteria. The highest value of -0.29 is obtained for a subject's own payoff and an equal distribution of all payoffs. However, according to a two-tailed t-test no coefficient turned out to be significant.

Table 7.3: Correlation coefficient between different criteria

N = 168	the amount of my payoff	the sum of all payoffs	difference between the lowest and highest payoff	equal distribution of payoffs
the amount of my payoff	1.00			
the sum of all payoffs	-0.09	1.00		
difference between the lowest and highest payoff	0.04	-0.03	1.00	
equal distribution of payoffs	-0.29	-0.14	-0.11	1.00

To summarize, most participants stated to have considered just one criterion when making their decision. For over 50% this was just their payoff. Thus, they neither focused to a large extent on the members of their delegation nor did they focus on all six players. However, at least batches of social behavior were indicated; a finding that holds for all decision-making specifications.

7.3 Qualitative content analysis

The questionnaire contained free-input fields for various questions. In these fields participants could describe with their own words how they reached their decisions. I used those answers to conduct a qualitative content analysis, which enabled me to use an inductive approach (Mayring, 2010, p. 67ff). This allowed me to identify important quotes and to generate categories out of the data itself. To discover these aspects I developed the coding guide (following Mayring, 2002), which is shown in Sec. A.16.[391] It contains different categories of quotes, their corresponding definition, example statements as well as coding rules. As according to the guide, the comments of the participants were allocated into the respective categories.

Tab. 7.4 shows the results of the respondent' comments in the free-input field "Please describe your decision rule with your own words?" Overall, 89% of participants replied to this question. Most subjects stated that they tried to maximize their own payoffs (30%) or their own payoffs under consideration of the other players' votes (43%). This corresponds to the results of the binary questionnaire (Tab. 7.1). Only a few participants were looking for a fair alternative (8%), and even less frequently they stated that their decision followed a Mini-max rule (3%) or was guided by a high sum off all points (3%).

Table 7.4: Free-input field "Please describe your decision rule with your own words?"

N = 168	
empty / no response	11%
"my own payoffs taking the probable vote of the other players into account"	43%
"my own payoffs"	30%
"fair / consensus alternative, none player worst"	8%
"minimize my risk"	3%
"highest average / sum off all points"	3%

Next, I asked which other players or payoffs the subjects focused on. Tab. 7.5 contains the results. 18% of all participants did not answer this question. The others focused, not surprisingly, mainly on themselves (41%). Looking at the other categories one interesting pattern can be seen. A subject's own delegation did not get more attention than both delegations (each 14%). About 7% focused on the payoffs of the other delegation. About 5% took into account which other players would be likely to vote for the same alternative. This corresponds to the decision rule mentioned above about "taking the probable vote of the other players into account". It is important to note that the evaluation of other players does not necessarily correspond with social preferences. The consideration of

[391] The coding guide in Sec. A.16 is only available in German.

other player's payoffs may be a first step when taking their well-being into account. In other words, it is a necessary but not a sufficient condition. Assessing the other payoffs is also perfectly in line with evaluating the probability of an alternative to get selected from a pure self-interested perspective.

Table 7.5: Free-input field "Which players did you focus on?"

N = 168	
empty / no response	18%
"myself"	41%
"my delegation"	14%
"the other delegation"	7%
"both delegations / all six players"	14%
"players (in both delegations) that are likely to vote with me"	5%

The next question investigated the criteria for making a decision. Tab. 7.6 shows that 14% of the subjects did not reply, and that for the rest the main focus rests on a subject's own payoff one more time. This is expressed in the amount of points (36%), by reaching at least the second-best result (4%) or by avoiding poor results (12%). Again, a considerable amount of statements mentioned the own payoff conditioned on the probability of success of the corresponding alternative (13%). The same logic underlies an alternative which poses at least four high payoffs (8%) as the threshold for an alternative to be selected as final outcome was at four players. Altruistic motives played only a minor role, the difference between the lowest and highest payoff is mentioned by 7% and an equal distribution of payoffs by 6% of the subjects.

Table 7.6: Free-input field "Which criterion of an alternative did you consider?"

N = 168	
empty / no response	14%
"the amount of my payoff"	36%
"avoid poor payoffs for myself"	12%
"at least my second-best payoff"	4%
"my own payoff, taking the probability of each alternative into account"	13%
"at least four high payoffs"	8%
"the difference between the lowest and highest payoff"	7%
"an equal distribution of the payoffs"	6%

The last free-input field asked the participants directly what aspect of their decision-making changed when the decision-making split up. This happened in different ways (cf. Sec. 4.2.6): Firstly, when the experiment moved from POL to SIM the allocation of voting competences split up. Secondly, the voting sequence was modified when altering between SIM and SEQ. Thirdly, both aspects changed when subjects took part in POL and SEQ.

Tab. 7.7 reveals that 38% of the participants left this text field empty and for 7% it made no difference. Taking a look at the respondents, 23% said that their own payoffs mattered more. The tactical aspect of the decision-making got more important for 14%. In the statements the subjects often indicated that they "tried" to anticipate the decision of other players. This suggests that they do not think of themselves as being successful in using this strategy. Furthermore, reaching a compromise (9%) and a subject's own delegation (7%) became more important.

Table 7.7: Free-input field "What changed when the decision-making procedure split up?"

N = 168	
empty / no response	38%
"nothing changed"	7%
"my own payoff mattered more"	23%
"it was more important to anticipated the decision of other players"	14%
"reaching a compromise was more important"	9%
"my own delegation got more important"	7%

Next, Tab. 7.8 shows the results of the same question separately, according to the different modifications of the decision-making procedure; i.e., which aspects changed between the first and second decision rule of an experimental session. Looking at the answers for the single modifications three interesting divergences emerge. Firstly, when altering the voting sequence fewer subjects stated that their own payoff had gotten more decisive (14% vs. 24% or 25%). Secondly, at the same time the anticipating of other players' votes became more important (31% vs. 13% or 17%). This suggests that introducing sequence into the decision mechanism led to more strategic voting. Thirdly, when dividing voting competences, reaching a compromise got more significant (14% vs. 3% or 5%). Here, the loss of voting rights seems to influence in favor of a more balanced result. Sec. 7.4 looks deeper into these differences.

Table 7.8: Free-input field "What changed when the decision-making procedure split up?" separated by treatment

Aspect of the decision-making procedure which was modified N	voting competence 72	voting sequence 36	voting competence and sequence 60
empty	35%	42%	40%
"nothing changed"	7%	6%	8%
"my own payoff mattered more"	24%	14%	25%
"it was more important to anticipated the decision of other players"	13%	31%	17%
"reaching a compromise was more important"	14%	3%	5%
"my own delegation got more important"	8%	6%	5%

Overall, the comments of the participants in all free-input fields confirmed that a subject's own payoff was the main driving force of decision-making. This insight is complemented by tactical considerations. Approximately 41% of the subjects described their decision rule as focused on their own payoffs, but they were "taking the probable vote of the other players into account". This approach adjusted each alternative by its specific cue validity; a strategy that became more important the more complex or uncertain the task encountered was.

7.4 Word scaling

In this section I use the same free-input fields as for the qualitative content analysis. The text was processed by the word scaling algorithm WORDFISH (Proksch and Slapin, 2009). This computational content analysis focused on the identification of similarities (respectively differences) from a subject's perspective when the decision rules changed. How did the participants describe their reaction to a modification of voting competences or sequence? In contrast to the analysis of the voting data, which is explicitly set in contrast to previous experiments (e.g., S&K, 2010), this investigation follows an inductive research interest.

WORDFISH is written in R statistical language (Version 2.14.2., release date 2012-02-19).[392] The algorithm extracts (spatial) positions from text documents. More precisely, using word frequencies WORDFISH places different documents into a single latent dimension of discrepancy by means of maximum likelihood estimations. The main argument for using WORDFISH was that it is a scaling technique and does not need any anchoring document. This distinguishes it from other computer-based content analysis approaches which use reference texts (e.g., Laver et al., 2003).[393] Instead, WORDFISH relies on a statistical model of word counts, more precisely a Poisson distribution (Poisson, 1837) of word frequencies. This proceeding fits my research questions as I did not have a specific baseline (or reference or anchor) document.

My analysis followed to a great extent the well documented example of Slapin and Proksch (2008). In a first preparation step the text data was "cleared up". I performed a spell check, removed quotation marks, hyphens, slash and backslash signs, etc. In addition, I had to replace the German specific mutated vowels as well as the letter "ß". Thus, the spelling of words became consistent across all documents.

Next, I applied the text mining package included in R (tm package, Feinerer et al., 2008). This package transforms all letters into lower case and removes all numbers from the text. In addition, it comprises a stemming algorithm and a stop-word dictionary. A stemmer algorithm omits morphological and inflexional endings from words and returns the stemmed root word, so similar words are captured as one. For example, a stemmer reduces the words "fishing", "fisher", and "fished" to the joint root word "fish". By using a stop-word dictionary the text is further cleared of all meaningless conjunctions, articles, etc. The package already contains dictionaries for many languages which can be edited manually. Applying these functions left a term-document matrix which had significantly fewer unique words than the original texts. This enabled a more efficient estimation.

[392] The homepage of the r-project can be found at `http://www.r-project.org`.

[393] For a comprehensive overview on different text-analysis methods cf. Slapin and Proksch (2008, p. 705-708) as well as Benoit et al. (2005) and Benoit and Laver (2007a).

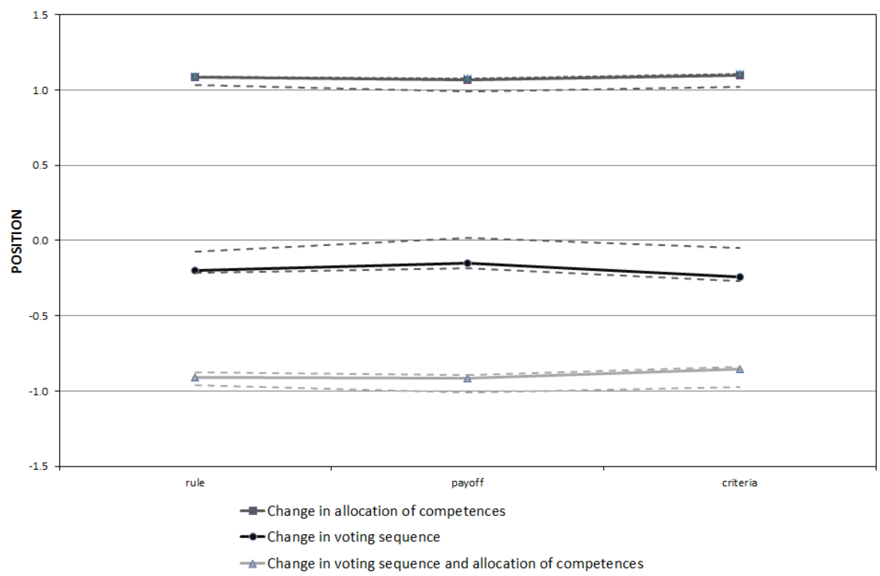
7.4.1 The latent dimension of discrepancy

When applying the WORDFISH algorithm to the documents the texts were structured in the following way. Firstly, I separated the comments into the categories decision rule (“Please describe your decision rule with your own words?”), focus of attention (“Which players or payoffs did you focus on?”) and decision criteria (“Which criteria of an alternative did you consider?”). These categories are from now on referred to as *rule*, *payoff* and *criteria*. Secondly, I divided the comments accordingly to the changes of procedure. Therefore, I was able to compare the reaction of participants when changing the allocation of competences and the voting sequence (alternate between POL and SEQ) to reactions when changing the allocation of competences (alternate between POL and SIM) or the voting sequence (alternate between SIM and SEQ) alone. Fig. 7.1 plots the position estimates of the latent discrepancy dimension.

Figure 7.1: WORDFISH: latent discrepancy dimension

EXPLANATORY NOTE

The figure depicts the WORDFISH estimates for the latent discrepancy dimension over three categories of comments. The chosen order of the categories is not meaningful but arbitrary. The dashed lines resemble 95% CI.



The absolute position values are not decisive; rather, the main results arise out of the relative spatial positions. Subjects experienced a change in allocation of competences alone as considerably different from a change in voting sequence or both modifications together. I found this pattern in all three categories of comments: whether I asked for decision rules, the focus of attention or the criteria the subjects took into account. The results were also validated by the obtained CI, which enabled a distinct demarcation.[394] Overall, the modification of competences or voting sequence resulted in significantly different reactions. When both modifications took place together, the effects of altering the voting sequence dominated and the reactions appeared even stronger.

[394] Following Slapin and Proksch (2008, p. 710), the CI were obtained through 500 times parametric bootstrapping (cf. Efron and Tibshirani, 1993). As uncertainty estimates they can be seen as robustness indicators.

7.4.2 Word weights and fixed-effects

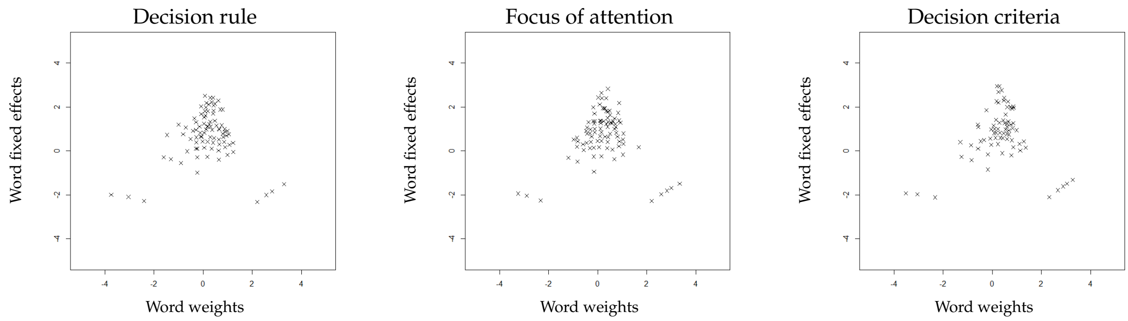
WORDFISH also offers the possibility to identify the words which substantiate the spatial positioning. The algorithm returns two parameters for every word. The first is a word fixed effect, which captures the fact that some words are used more frequently than others in all documents. The second displays a word's weight. This resembles the importance of each word in discriminating between positions. An effective analysis should result in two types of word categories: i) frequent words without spatial meaning, resulting in large fixed effects and low weights, and ii) substantial words for the spatial discrimination, containing lower fixed effects but larger word weights. A corresponding two-dimensional scatter-plot should reflect these findings by resembling an "Eiffel tower of words" (a slender peak and a broad basement with a blank center, Slapin and Proksch, 2008, p. 715).

The scatter-plots in Fig. 7.2 contain the estimated word fixed effects and word weights for the three comment categories. Importantly, the labeling of positive and negative effects or weights only facilitates the interpretation of the plot. Those are just spatial categories calculated by the WORDFISH algorithm, representing left and right or up and down. The terms have no pejorative or favorable meaning. Furthermore, no category is preferable to the other. Looking at the figures the expectations are confirmed. Words with high fixed effects scatter around a weight of zero, but words with low fixed effects show a higher absolute weight. The patterns are the same for all three categories.

Figure 7.2: WORDFISH: word-weights vs. word fixed effects

EXPLANATORY NOTE

The figures depict the WORDFISH estimates for word fixed effects and word weights. The results are separated according to the three free-input fields of decision rule, focus of attention and decision criteria.



Tab. 7.9 lists the words with the highest impact factor in alphabetical order.[395] Again, the labels exhibit no judging or quality relation of the two directions. Looking at the decision rules, positive comments mentioned frequently terms as "balanced", "enforceable", "fair", "majorities" or "four players". Negative comments contained more often such terms as "individual", "utility maximizing" or "profit". This indicated that the spatial dimension resembles a kind of universalism - egoism dimension.

[395] As robustness test I conducted the calculations with and without applying the stemming algorithm to the text documents. The results do not change between these two modifications.

Table 7.9: Words with the highest impact factor

Category		Terms	Classification
Decision rule	positive	balance, compare, enforceable, fair, four players, good, higher, largest, majorities, maximum, one, optimal, part, strategy	universalism
	negative	area, field, high, individual, little, profit, search, sum, utility maximizing	egoism
Focus of attention	positive	agree, at least, fifty points, game, good, large, most likely, oriented, others, second-best, solution, three players, twenty points, unlikely	exclusive
	negative	aware, column selection, field, high, largely, low, overall, possible solution, respond, row selection, six players, worst	inclusive
Decision criteria	positive	adjust, average, equal, few players, group, more players, much, opportunity, payoffs, prefer, profit, realistic, simultaneously	moderate
	negative	all players, difference, majority, never, one alternative, points, probably, strategy, delegation members	strategic utility maximization

Taking a look at the comments on the focus of attention the terms “agree”, “others” and “three players” were in the positive part. The negative side contained often the expressions “overall”, “column selection”, “row selection” and “six players”. This indicated an inclusive - exclusive aspect (whole group vs. own delegation). Both sides contained references to the strategic aspect of the decision problem (“unlikely”, “most likely” or “possible solution”) and a resulting risk avoidance (“second-best” or “worst”).

The positive comments on decision-making criteria were “equal”, “group”, “average”, “more players” and “realistic”. On the negative side were the terms “all players”, “difference”, “majority”, “probably” and “strategy”. So the positive aspects were categorized as moderate and consensus-orientated and the negative terms pointed towards strategic utility-maximization.

To illustrate, Tab. 7.10 shows words with a word weight around zero. These words were mentioned very often but did not exert a significant influence when calculating positions. This does not mean that these terms were irrelevant when looking at the decision-making behavior of the participants in the experiment. The low weights just indicate that they were not used differently between the procedures. They were often used under all decision rules! It may be that these phrases resembled basic and important aspects of the decision-making process in general. They were not affected by different forms of voting competences or decision sequences.

Referring to the results in Tab. 7.9 the left (or negative) dimension contained the aspects egoism, inclusiveness and strategic utility maximization. The right (or positive) dimension, on the other hand, referred to the terms universalism, exclusiveness and moderate behavior. Linking this to the findings of the spatial position estimations Fig. 7.1 I was able to associate the decision rule modifications with corresponding behavioral patterns.

Table 7.10: None decisive terms

Category	Terms
Decision rule	agreement, alternative, anticipate, arithmetic, balanced, choose, comparison, compromise, consequence, consideration, correlation, criterion, doubt, estimate, generality, group members, maximizes, maximum value, members, opportunity, payoff, players, profit, similar, simultaneously, stable, successfully, sure, delegation, uncertain, vote, winning chance
Focus of attention	anticipate, attempt, avoid, balanced, beginning, best, both, choice for developing, circumstances, column, combination, compromises, concentrated, consequence, decision, error, fragrance, initially, intuitive, largest, low, lowest, maximum, necessarily, odds, options, payoffs, players, points, predict, probability, probable, profit, relatively, result, round, satisfied, scheme, significantly, suspected, delegation
Decision criteria	accordingly, alternately, alternative, arithmetic, at least, balanced, best, center, difficult, eliminate, euro, experiment, failure, high, interest, loss, lowest, maximum, minimized, more, mostly, optimal, own, plus, potential, predict, prefer, purchase, safe, small, superior, highest, variance, variation, yield

A change of the allocation of competences (between POL and SIM) was associated with positive positions. This linked voting competences to thoughts of universalism and exclusiveness. Participants considered who had to be taken into account for reaching a consensus. The moderate behavior may also indicate a risk minimizing approach. By giving up the overall highest payoff a subject aimed to avoid the worst possibilities.

A change in voting sequence (between SIM and SEQ) or modifying both aspects together (between POL and SEQ) received negative position estimates. This linked the voting sequence to egoism and strategic utility maximization behavior. Thus, subjects seem to have captured the theoretical implication of an alternate solution concept.

7.5 Chapter summary

This chapter supplements the empirical analysis of the experiment. In addition to the voting data it contributes insights from the participants' perspective. The post-experiment survey allowed for general comments on the experiment but focused on the decision-making process of the participants. It contained 'yes or no' questions as well as free-input fields.

The general comments proved the desire of subjects to communicate. Across all procedures the collective task sparked the desire for consultation opportunities. This desire was expressed as a wish, not as a complaint. Some comments criticized the duration of and delays within the sessions, but delays cannot be avoided in collective tasks. Most importantly, I found no evidence for disturbing effects such as EDE.

The binary questionnaire revealed that most participants considered just one criterion when making their decision, and for about 50% this was, not surprisingly, just their own payoff. In addition, at least a few social considerations were indicated, a finding that holds for all decision-making specifications.

The free-input fields were analyzed using a qualitative content analysis as well as a computational text analysis. The qualitative content analysis confirmed the binary findings. A subject's own payoff was the main driving force for the decisions made in the experiment. The free-input fields supplemented this behavior with a tactical aspect as subjects took "the probable vote of the other players into account". Thus, the periphrastic expression of sophisticated voting receives a more vivid description. Such a procedure was used with increasing frequency, the more complex or uncertain the decision-making procedure got.

The computational text analysis indicated that the alternation of the decision-making procedure influenced the motives of the participants. A change of decision-making competences immediately triggered the participants' thinking patterns for who had to be taken into account. This led to a higher exclusiveness in voting. A change in voting sequence was processed through adjusted strategic utility maximization. Both patterns are in line with the qualitative content analysis as well as the binary questions.

Overall, the most important insight of the post-experiment survey was its validation of the design. Subjects did not misunderstand the experimental instructions. This would have led to dissimilarities between the perceptions of what a subject thought they were doing and their actual behavior. Fortunately, this was not observed.

8 Complex reality and limited models

The previous parts of my study plainly show the relevance of NSP; I demonstrate the vulnerability of inference when falsely assuming separable preferences (cf. Chap. 3) and clarify the impact of nonseparability on collective (cf. Chap. 5) as well as on individual (cf. Chap. 6) decision-making. Yet, I focus only on the knowledge gained if NSP are included into analytical research. The corresponding requirements for the consideration of nonseparability have so far only been mentioned in passing. In this chapter I rectify this deficiency and add the missing assessment of the necessary expenses for including NSP.

In Sec. 8.1 I discuss possible empirical implementations and offer a rough estimation of related cost in terms of necessary tasks and methods. This enables an assessment of the knowledge gained with respect to the required effort of modeling. The subsequent sections situate the argument in a broader context. All research must decide between including more details and keeping models feasible. Sec. 8.2 discusses the concept of abstraction, its application in analytical models and its legitimacy. Hereby, I set the identified threat to inference (ARGUMENT 1) against the benefits of simplifying assumptions. Subsequently, I trace in Sec. 8.3 the latest development in social science research paradigms in which a 'new thinking' suggests a shift away from artificial assumptions towards a more realistic modeling of behavior. This affects analytical research in general and is most certainly relevant for the assumption of separability. My experiment clearly demonstrated the behavioral effects of NSP (ARGUMENT 2 and ARGUMENT 3), which therefore should not be ignored. Finally, Sec. 8.4 summarizes whether and how nonseparability should be included into analytical research. It also places my contribution into the current state of political science.

8.1 The necessary effort

The previous chapters investigate the possible gains of including NSP into scientific research. I use both field and laboratory data for this assessment. What stands against their observance are the necessary requirements for including nonseparability, which are twofold. In either case, the corresponding data has to be i) collected and ii) incorporated into the analysis. Both tasks make claims that exceed those of standard approaches and which might challenge researchers.

In the first step the necessary data must be gathered. Collecting data can turn into an exhausting experience, and more data most certainly implies more cost. It is obvious that collecting data on NSP will lead to higher expenditures. First of all, this concerns purely monetary resources for access rights, licenses, etc. Then, of course, also supplemental work hours and additional staff might be necessary. The coding of NSP in the DEU data set (Thomson et al., 2006), reported in Sec. 3.4, may serve as an example. At least three graduate student coders were necessary to obtain sufficient reliability. All of them had to receive a thorough introduction to the concept of nonseparability and spent quite a few hours of work on the coding of the 66 law proposals. Yet, such additional effort is inevitable for an appropriate data provision. The following paragraphs provide an overview on such tasks by summarizing the pros and cons of different methods. Unfortunately, none of them comes without problems.[396]

SURVEYS AND INTERVIEWS

Like other data on individual decision-making (e.g., ideal points, salience and believes), the information on NSP can be collected through interviews and surveys. While the measurement of separable preferences can be limited to an actors' first preference, NSP require an evaluation of actors' utility function at several values. Thus, the collection turns into a disproportionately more complex task (cf. Sec. 2.4).

An accurate example for measuring NSP can be found in Lacy (2001b). The author measured NSP in survey respondents' preferences for the U.S.[397] In his survey he used the following arrangement of questions to identify NSP (Lacy, 2001b, p. 253ff). First, respondents were asked for their unconditional first preference on two issues. Second, in order to identify NSP, the issues were alternately fixed at specific levels in follow-up questions which asked for the resulting modification of the other issue. This questioning was repeated for all combinations and orders of issues. The study clearly showed that common survey questions about ideal position are insufficient when dealing with NSP (cf. Sec. 2.4.6). The same proceeding holds true for expert interviews. The criteria for selecting experts stay the same as when only concerned with separable data. But in the proceeding of the interview the questions have to be structured in the same manner as used by Lacy (2001b).

It holds for both data-collecting techniques that the hypothetical character of the follow-up questions is demanding. It represents a challenge to the imagination as well as the attention of the respondent and stresses its patience. The additional hypothetical evaluations of actors' utility functions at several values leads to longer interviews, much more questions and a more complex task.

In addition, questions repeated iteratively may develop distorting self-dynamics. If one asks people many times about hypothetical scenarios this will affect the answers given.

[396] Cf. Schnapp et al. (2009) for a comprehensive overview on the collection and types of data commonly used in political science research.

[397] Lacy (2001a,b) found that preferences on income taxation depended on crime prevention policies, preferences on environmental pollution depended on environmental regulation, preferences on defense spending depended on social spending, preferences on immigration policy depended on the constitutional status of English being the only official language in the U.S., etc.

For example, in such interviews politicians do not want to look like a pushover. So they announce to stick to their current policy (e.g., to liberalize certain industries), although they would prefer an adjustment under the scenarios discussed (e.g., altered regulatory safeguards).

OBSERVATIONAL DATA

In addition to directly asking individuals for their preferences, they can also be obtained through behavioral observations. Such observations are not restricted to a specific setting; instead, they can take place in a subject's everyday environment. This noninterference enables surveyors to observe a subject's natural behavior without distortion (cf. Ortmann, 2005). Observational data generally offers larger amounts of data, especially in the form of longitudinal data which is otherwise difficult to obtain. However, establishing causal effects can be statistically challenging (e.g., Hill et al., 2005).

As discussed in Sec. 6.2, two types of behavior are commonly observed in electoral processes; these are sincere as well as strategic voting (Herrmann, 2012). When aiming to assess their respective spreading, the problem is that "to identify strategic voting requires that we know both the voter's true values and the voter's actual expression of the values in a vote. From direct observation we can know only the latter. We must infer the former from other and softer evidence" (Riker, 1982a, p. 167). This makes it difficult to distinguish between them. "If we rely only on the observable actions (e.g., votes) of legislators to test theories of strategic behavior we cannot determine whether the observed behavior represents legislators' true preferences or whether the legislators are acting strategically" (Clinton and Meirowitz, 2004, p. 676).[398]

To uncover an individual's true preferences is also of paramount importance for identifying nonseparability. Yet, in particular Chap. 3 reveals in detail how much effort it takes to obtain individual data. NSP are subject to the same limitations and it seems difficult to uncover the relevant relations using solely observational data.[399] Observed behavior nearly always offers "multiple causal paths" (Braumöller, 2003, p. 209). This implies a high degree of "substitutability" (Braumöller, 2003, p. 215) in explanation of the data which can resemble nothing more than an educated guess for the truth.[400]

Observational data may often provide in depth knowledge on a specific effect but not on universal patterns. The argument made by Levitt and List (2007, p. 160) concerning low "cross-situational consistency of behavior" describes this very well. As long as NSP are not well studied and the relevant relationships and policies are not known, this ap-

[398] While empirical investigations of legislative strategic voting are relatively scarce, cf. Clinton and Meirowitz (2004) for an overview of such approaches.

[399] I am not arguing that investigating nonseparability with observational data is impossible. In fact, the problems encountered are oftentimes just similar compared to conducting a survey (e.g., costs and opportunities of data collection). Yet, the previous permanent neglect of NSP does not provide the necessary empirical framework.

[400] Braumöller (2003, p. 215, table 1) offered a selection of examples for international relations, comparative and American politics.

proach is indeterminate but might be worthwhile in future research. Therefore, I discuss in Sec. 9.2.2 the possibility of field experiments, i.e., a first step into this direction.

EXPERIMENTS

With respect to validity, laboratory experiments are the opposite of observational research. The pros and cons of this method are assessed in Sec. 4.1 in detail. In short, experiments possess a high degree of internal validity at the expense of external validity (Schram, 2005). Inducing preferences according to the research question is only possible in the laboratory (McDermott, 2002, p. 326). In particular, the monitoring capabilities as well as the likelihood of eliminating interfering effects are excellent.

Of course, this tells us nothing about where NSP exist in reality. Such "proof of external validity is always empirical" (Morton and Williams, 2010, p. 196). That is why further experiments and other complementary studies are inevitable. Yet, an experiment enables the investigation of the effects which NSP exert on decision-making processes; thus, for "theory development, testing and refinement" (McDermott, 2002, p. 341). In other words, one can learn about nonseparability without a long and onerous search for suitable situations. Of course, such investigations can only be the first step in a series of further analyses; but that is the exact purpose of an experiment and it fulfills it pretty well (cf. Sec. 1.3.2).

SIMULATIONS

The previous methods aim to actually measure NSP. In Chap. 3 I apply a different approach; the determination of NSP is accomplished by means of simulation techniques. This was the only possible way as the original data collection (through expert interviews) was not concerned with nonseparability. This leaves the technique with the limitation that is it only possible to theorize on and not prove the existence of NSP.

The simulation was based on three steps with each focusing on a single aspect of NSP (cf. Sec. 3.4): existence, direction and reciprocity. Firstly, it was identified whether or not actors' preferences over two issues may be nonseparable at all. Secondly, the direction of the nonseparability was added, i.e., whether the issues in question are supplements or complements. Third, it was determined whether the potential nonseparability is reciprocal or not (cf. Sec. 2.4.5). This coding scheme was then used for a computational comparison of legislative models of decision-making.

Simulations should only be used in combination with a model or theory already well proven (Clarke and Primo, 2007). The unconstrained bargaining model, constrained bargaining model and agenda-setting model used in Chap. 3 fulfill this claim as they can rely on a vast literature on legislative of decision-making (cf. Sec. 3.1).

To summarize, Tab. 8.1 lists the pros and cons of the different methods. As neither of them comes without flaws, the most promising way forward might be to combine the various approaches. Together they can balance each other's respective problems, as their strengths are complementary. For example, Hamenstädt (2012) argued in favor off bringing the lab into the field in order to combine the explanatory powers of both methods. Sec. 9.2.2 will illustrate this principle further and demonstrate its possible significance for future research. Of course, this is no easy task, but the numerous opportunities for scientific analysis allow a closer examination of the nonseparability phenomenon.

Table 8.1: Methods of data collection for investigating nonseparability

EXPLANATORY NOTE
The table lists possible methods of data collection and their pros and cons with respect to the assembling of the necessary information to determine NSP. To a large extent these points agree with general characterizations of the methods (e.g., Moses and Knutsen, 2007).

METHOD	Pro	Contra
Surveys and (expert) interviews	determine NSP exactly	hypothetical questions and long sessions necessary
Observational data	obtain "true" NSP from a natural setting	multiple possible explanations make a determination difficult
Laboratory experiments	induce preferences according to research question	reaction to induced NSP must not be the same outside the laboratory
Simulations	feasible even ex-post	applicable only with well proven models; can only theorize on and not prove the existence of NSP

EMPIRICAL EVALUATION

The second step of incorporating the eventually collected data comprises the operationalization of NSP. Here, Sec. 2.4 discusses the theoretical aspects in detail and Chap. 3 supplements an empirical implementation into different legislative models. In general, models become more complex and demanding as they incorporate additional information. However, the fact that a problem is somewhat more complicated alone is no reason to refrain from scientific investigation. Improvements in computational technology enable the calculation of most sophisticated theories (cf. Humphreys, 2004).

Overall, the claims represented to the researcher when setting up the model including NSP are not unreasonably high. Accuracy is an inevitable requirement to which scientific research must comply. This holds also for the assessment of nonseparability. The (admittedly) high demands on data quality and computational resources are no valid counter-argument.

8.2 The concept of abstraction

Abstraction resembles a fundamental aspect of analytical research, especially when specifying a theoretical model. Reality is too complex to comprehend without simplification (Kesten and Pnueli, 1998). As the only completely accurate way to map a city is to set up a

1:1 template, even a vast map with a gigantic resolution would omit data. Weber argued that it is not important if a model does not mirror reality perfectly and in every detail (cf. Weber, 1968, p. 190ff).[401] In fact, he stated that "ideal types" (e.g., a perfectly and rigorously rational voter) only at all exist in abstract concepts (cf. Shils and Finch, 1949, p. 89-95). Yet, without a model it would not be possible to gauge real behavior. As early as 1836, John Stuart Mill's famous essay "On the Definition of Political Economy; and on the Method of Investigation Proper to It" took advantage of "a hypothetical subject, whose narrow and well-defined motives made him a useful abstraction in economic analysis" (Persky, 1995, p. 222-223). Miljkovic (2012) aptly summarized in her preamble that the "complexity of nature and consequently the complexity of everyday life processes often make the mathematical models deterministically unsolvable. Moreover, if such solutions do exist, usually a lot of resources are required to find them. Therefore, the idea of approximation has developed as an irreplaceable tool for handling many problems."

A model resembles a conceptual representation of the real data generating process (Morton and Williams, 2010, p. 194). Yet, by its very definition a model is restricted to some aspects of a real phenomenon and different models can focus on completely different aspects of the same event (Lave and March, 1975, p. 3ff).[402] To correctly comprehend the model, its implications and predictions requires knowing its original purpose and the assumptions made when setting it up (cf. Grosslight et al., 1991). The performance assessment of every model depends on multiple characteristics, such as validity, reliability and consistency. Downs and Wildenmann (1968) advocated that theoretical models should first of all be evaluated with respect to their predictive accuracy and not the reference to reality of their assumptions. However, a model's 'cost of use', especially in comparison to alternative models, is essential.

Each systematic approach must thoroughly consider the assumptions made during the research process. Bloomfield and Anderson (2010) stated that it is necessary to think more clearly about the nature of assumptions in their discipline of economics. Bloomfield et al. (2009) discussed typical categories of assumptions used by experimentalists. They identified structural assumptions which describe the institutions in which agents interact (e.g., information distribution, possible set of actions and incentives), behavioral assumptions which characterize agents' preferences and decision-making (e.g., the form of the utility function) and equilibrium assumptions that describe the solution concepts applied (e.g., Bayesian Nash equilibrium or backward induction). These categories must all be considered when setting up an analytical model.

[401] Klein (1985) examined the mode (proximity or structure) and dimensions (physical, temporal, attributional or construal) in which abstraction takes place. He identified a limit for the principle of abstraction "since people cannot handle substantial amounts of construal abstraction" (Klein, 1985, p. 677) (which represents a participant's subjective perception of the content of a model). Here, "anything more than trivial abstraction in this dimension alters the meaning of models to such an extent that human players will not be able properly to understand them" (Klein, 1985, p. 671).

[402] Long (2006, p. 5) clarified that in the very Aristotelian tradition abstraction is to be understood "as a matter of attending to some aspects of a thing and ignoring others."

In general, each additional assumption complicates the creation of a model. So why not stay with a few assumptions and slim models? Looking into the field of behavioral economics, and particularly into research on social preferences, Mertins (2008, p. 33) pointed out that a "large literature has shown that we can go a surprisingly long way with very simple models of fairness in some important classes of games." Thus, why should more and more complex and intricate models be necessary at all? This depends on the knowledge gained (whether it is worth the effort) and the reliability of the findings. Cameron and Morton (2002, p. 793) explained that "even inadequate formal models" possess virtue as they "have the advantage of clarity."

Any concepts with substantial impact should be included into prospective scientific research.[403] Otherwise a simplifying assumption (in case of my research the presumption of separable preferences) has its legitimacy. This follows Occam's razor[404] which calls, all else equal, for the simplest model. Occam's razor constitutes a widely used principle in economics.[405] Yet while this norm has provided valuable assistance for scientific research, "its continued use [...] risks significant opportunities to be missed" (Domingos, 1999, p. 409). Reid (1987, p. 551) warned that a "too ready application of Occam's razor' (broadly defined to champion elegance and simplicity) slows rather than speeds the growth of economic knowledge." There exists even stronger criticism[406] which is part of an intense controversy about its application in science in general.[407] Thus, a thorough investigation of possible insights is inevitable before rejecting alternative specifications.

This discussion, so far, seems to refer mostly to the discipline of economics. Yet, all analytical research has to face an assessment of more details included and manageable limits of modeling. This also holds true for political science, even though a famous quote of Otto von Bismarck states that "Politics is not an exact science" (Speech to the Herrenhaus, Otto von Bismarck, 1863). Aptly enough, Noel (2010) replied that "politics is not a science, but it can be studied systematically" and thus be subject to analytical research.

From the beginning of this study it was laid out that researchers have to face a trade-off between excluding NSP and keeping the analysis simple, or including NSP and dealing

[403] In addition to adding new findings, confirming or refuting of existing knowledge is also a valuable contribution.

[404] The term goes back to William of Occam who stated in the late middle ages that "nunquam ponenda est pluralitas sin necesitate" (Domingos, 1999, p. 409). This was translated by Tornay (1938) as "entities should not be multiplied beyond necessity".

[405] In his marvelous description of the history of economic ideas, Heilbronner (2000, p. 103) showed that it was David Ricardo who "gave the powerful tool of abstraction to economics."

[406] An aptly example of such criticism is the contribution "Razoring Ockham's razor" (`http://rationallyspeaking.blogspot.de/2011/05/razoring-ockhams-razor.html`, posted at May 6^{th} 2011) by Pigliucci (2013) where the author emphasized the necessity for a discussion on the principle's proper application. He also described a contradiction within Ockham's razor; namely that "philosophers often refer to this as the principle of economy, while scientists tend to call it parsimony. Skeptics invoke it every time they wish to dismiss out of hand claims of unusual phenomena (after all, to invoke the "unusual" is by definition unparsimonious, so there)."

[407] Riesch (2010) interviewed 40 scientists on their views of Occam's razor and simplicity. She received various interpretations of the principle; the responses ranged from complete rejection to the assessment that "Occam's razor indeed forms an integral part of scientific method" (Riesch, 2010, p. 86).

with more complex requirements. A complex reality (including NSP) and abstracting models (frequently ignoring NSP) are difficult to reconcile. So far nonseparability has often been neglected. The justifications for not taking account of NSP varied. For "the purpose of simplifying" (e.g., Morgan, 1990, p. 321) as well as because it "is common" (e.g., Le Breton and Sen, 1999, p. 606) were the most frequent arguments.

I do not question the general principle of abstraction, only its overuse. The threat neglected nonseparability poses to the inference of analytical models clearly substantiates the danger of overexploitation. Again, this is not a question which applies only to the topic of NSP. Reflecting on the adequacy of abstraction in economics, the sociologist Emil Durkheim made the noteworthy remark as early as in 1887 that "without doubt, in the field of economics the application of abstraction is legitimate" (Emil Durkheim, as quoted by Swedberg and Maurer, 2009, p. 53). But he also pointed out that "just not all abstractions are equally correct. Abstraction requires isolating a part of reality, not to let it disappear" (ibid.).

By ignoring NSP, fundamental aspects of decision-making are faded out of analytical consideration in favor of clearer mathematical models. Whether this is an appropriate way of investigating social phenomena is highly questionable. My study can only be the first step in considering this question thoroughly.

8.3 A more realistic view of human behavior

In his noteworthy essay "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" the physicist Eugene Wigner (1960) pointed out that the mathematical structure of physics itself has again and again led to further discoveries. His main argument was "that mathematical concepts turn up in entirely unexpected connections. Moreover, they often permit an unexpectedly close and accurate description of the phenomena in these connections". This could not be a mere coincidence, even if we thus far "do not understand the reasons of their usefulness" (Wigner, 1960).

This work has inspired much research on the phenomenon that "the enormous usefulness of the same piece of mathematics in widely different situations has no rational explanation" (Hamming, 1980, p. 82). Maybe there is a deeper reason why (it seems) "that the laws of nature are written in the language of mathematics" (Galileo Galilei, as quoted by Hamming, 1980, p. 82). Over the last fifty years, many contributions across disciplines followed up by assessing the usefulness of mathematical applications for their own field of research (e.g., Lesk (2000) for molecular biology and Tegmark (2007) for physics).[408] The idea was not only appreciated but also criticized, e.g., Hamming (1980) and Gray (2011) spoke against Wigner (1960) for giving only partial explanations or staying vague

[408]For an up to date review of contributions cf. Russ (2011).

in his argumentation. Nevertheless, both also stated that "these problems are in fact deep questions and worthy of further investigation" (Russ, 2011, p. 211).

Winger's contribution was written over 50 years ago; this raises the question if it is still a contemporary issue. With regard to its future relevance Djorgovski (2005, p. 131) argued that "applied computer science is now playing the role which mathematics did from the 17^{th} through the 20^{th} centuries: providing an orderly, formal framework and exploratory apparatus for other sciences." If so, there is a shift into a more technical dimension, but the basic phenomena that natural science is able to express even the most complicated relations with simple formulas remains.

THE DIFFERENCE BETWEEN SOCIAL AND NATURAL SCIENCES

The ability of natural science to describe their object of study with simple equations has triggered a desire in social science to work in the same way. However, we reach limits when aiming to describe humans and their behavior in simple formulas. This is most evident when "economists suffer from physics envy over their inability to neatly model human behavior" (Halevy et al., 2009, p. 8). Velupillai (2005, p. 849) even claimed "that mathematical economics is unreasonably ineffective."

The difference between the disciplines has been summarized aptly by Taleb (2007). He concluded that "if you know all possible conditions of a physical system you can [...] project its behaviour into the future. But this only concerns inanimate objects. We hit a stumbling block when social matters are involved. It is another matter to project a future when humans are involved, if you consider them living beings and endowed with free will" (Taleb, 2007, p. 183). It is inevitable to accept that social dynamics do not follow incontrovertible laws of nature. North (2008, p. 25) emphasized that we exist in "a non-ergotic world. An ergotic world would be one in which the fundamental underlying structure is uniform and exists everywhere. In such a world, if you understand that fundamental underlying structure and you want to solve a new problem, you go back to fundamentals and then build your theory based on the structure. Now that is what is done in the physical sciences and the natural sciences. The social sciences, however, have no such tools; and, what is much more difficult - the world just keeps changing."

This leaves the question what opportunities social science has left, if the focus on mathematical applications provides no answer. All in all, there are two possibilities. On one hand, research can aim for an alternative tool which serves as simplification device. On the other hand, research can engage and incorporate complexity as good as possible.

Halevy et al. (2009) advocated "The Unreasonable Effectiveness of Data". The authors focused mainly on their discipline of data mining and computer technology. Yet, their argument is valid for all scientific research[409] when applying optimization algorithms,

[409] Already Achen (1983) discussed the amount of data collected for political science research. While he advocated a development "towards theories of data", he also pointed to risks in this undertaking; namely measurement error and aggregation bias.

scrutinizing large amounts of data or formulating extensive systems of equations. The improvements in computer technology and the enormous growth of available data[410] enable the investigation of previously inaccessible research questions (cf. Humphreys, 2004). Here, self-learning algorithms that search through exabytes of data open many doors for science and industry.[411] Conducting laboratory experiments may represent an important building block for this prospective approach, as it enables researchers to obtain accurate information in a controlled environment.

As an alternative, Russ (2011, p. 211) promoted the necessity to accept complexity. He argued that focusing on rigid and idealized models or theories neither represents the present nor the future of social research. This calls for a less idealized but more realistic representation of behavior (Ariely, 2009) which takes into account the inherent complexity of social systems. For example, the intricate dependencies within human preferences.

THE INHERENT COMPLEXITY OF SOCIAL SYSTEMS

Political science was and is heavily influenced by neighboring scientific fields. Looking at its methodological framework Beck (2000) defined it as "welcoming discipline". The author argued that political scientists "use a variety of methods to attack questions related to political institutions and behavior. Although the methodological issues are defined by our political questions, we freely use whatever methodological solutions are available. Thus political methodology has freely drawn on insights from econometrics, psychometrics, sociology, and statistic" (Beck, 2000, p. 651).[412] Accordingly, Druckman and Lupia (2006, p. 18) emphasized that "context, not methodology, is what unites our discipline" when assessing political science.

Not surprisingly, a large part of quantitative political research followed neoclassical economics to an assumption-centered research approach which is characterized by a "rigorous corset of formal and propositional logic" (Ruckriegel, 2010, p. 3).[413] Bradley (2006, p. 17) summarized the neoclassical paradigm, admittedly somewhat exaggerated, as following Pythagoras' doctrine that "all is number".[414] Of course, the formalization was not simply a goal in itself. The introduction of the Homo economicus followed the purpose to pursue social science in analogy to natural science tradition as an exact science (Matis, 2007). The intent of the highly formalized approach was to offer the "availability of new logical and mathematical tools" (Walliser, 2008, p. 2).

[410]The study of Manhart (2011) documented this growth in detail.

[411]The study of Velten and Janata (2012) investigated the "explosion of data availability" and how this changed not only the IT but all business sectors.

[412]This collecting of methods from other fields has also been criticized. Achen (1983, p. 70) argued that "techniques invented by statisticians, psychologist, and economics [were] often meant for very different tasks". A specific version of this general criticism is the disapproval of the common and uncritical use of Euclidean utility functions to model political decision-making (cf. Sec. 2.4.2) by Milyo (2000b,a) and Benoit and Laver (2007b).

[413]This was not without controversy and in some fields of political science also a counter-movement was visible (cf. Monroe, 2005).

[414]Cf. Gill (2006) for a brilliant introduction into mathematical concepts in social science.

However, some assumptions made when introducing the Homo economicus are highly controversial and have often been called into question (e.g., Opp, 1999; Taylor, 2006). The "optimization mathematics forced economists to make very ambitious assumptions about the intellectual capacity of its agents - the controversial assumption of perfect rationality" (Geisendorf, 2009, p. 163) is one of the most prominent examples. This doubt about the rationality assumption is also reflected in my experimental results. The complexity of decision situations affected by NSP clearly exemplified human cognitive limitations.[415] Also, a "false assumption is that almost all people, almost all of the time, make choices that are in their best interest or at the very least are better than the choices that would be made by someone else" (Thaler and Sunstein, 2009, p. 9). In particular, Ariely (2008, p. xii) provided "a wide range of scientific experiments, findings, and anecdotes that [exemplify] how systematic certain mistakes are." Thus, "orthodox conceptions of rationality are evidently internally deficient and inadequate for explaining human interaction" (Colman, 2003, p. 139). McFadden (2006, p. 10-11) concluded that "Homo economicus, sovereign in tastes, steely-eyed and point-on in perception of risk, and relentless in maximization of happiness, is a rare species."

The simplifying assumptions implied an artificially high level of (only seemingly) precision and scientific reliance (cf. Ortlieb, 2010). The method was bound to a serious internal contradiction. On the one hand, complex decision situations and complicated action portfolios were traced back to basic patterns based on simple models to analyze the dynamics and structures of (political) behavior (Ostrom, 2005). On the other hand, obtaining the desired parsimony of the model required a large amount of complexity reduction through omission of less important aspects defined beforehand (Braun, 2013, p. 182).

Today, this is no longer state of the art.[416] Mankiw and Taylor (2010, p. 864) pointed out that "economics may borrow some methodology from the hard sciences but as a science of human behavior some of these methods are built on ever shifting sand." Instead of developing even more sophisticated models, economics is following a development that might be described as "rediscovering the human side in economics - from neoclassic back(!) to behavioral economics" (Ruckriegel, 2010, p. 1) or "the return of the lost human - ways to Homo sapiens economicus" (Dopfner, 2002). It follows the insight that "modern economics must be based on a realistic description of human behavior - not, as previously, on the assumption that we all act rationally" (Krugman, 2010). This leads to "an economic theory based on the actual behavior of people, not [...] one that is based on how people should behave" (Ariely, 2008, p. 265).[417]

[415]This is evidenced by the prevailing position of sincere behavior, especially under SIM.

[416]For an overview on the development of the rational choice paradigm cf. Gilboa (2010).

[417]Fehr and Schwarz (2002) summarized this change incisive. For outsiders economics often seemed like 'rocket science'; like a highly technical juggling with data and formulas looking for (a relatively mechanistic understood) causality. This ignored that economics is ultimately a human science; i.e., a science of human behavior. Yet, this view has in the last few years become increasingly important again. The authors attributed an important role in this development to "to the advent of micro-economic experiments, which brought about almost a methodical revolution" (Fehr and Schwarz, 2002, p. 5).

The shortcomings of current theories and models are diverse.[418] Akerlof (2007) assessed macroeconomic theory based on the traditional neoclassic assumptions of the Homo economicus as leading to inaccurate conclusions. In his remarkable speech at MIT, Friedman (2007) introduced the term "gross-individual-product". He claimed that it constitutes a necessary adjustment of outdated economic theories which had not "fully been able to capture what is happening far below the firm level, on the individual level" (Friedman, 2007). In his view this resembles a mismatch between the individual and the collective. The standard research framework is too far away from what is observed in the economy today. A similar plea was made by Kaufmann (2012), who argued in favor of breaking out of standard assumptions and to pay attention to the ethical questions of competition, instead of following the mantra of the free-market indiscriminately.

This call for a more accurate economic theory was strengthened by the recent global financial crisis (Akerlof and Shiller, 2009). A purely rational framework was no longer sufficient (Ariely, 2009). The speculative bubbles turned the focus on irrational behavior and "animal spirits" in the markets (Keynes, 1936; Dow and Dow, 2011). The financial crisis was not accountable within the standard economic models because it would have been too complicated to calculate (Johnson, 2012). However, Dalio (2011)[419] provided a framework for economic forecast which differed from the traditional perspective on supply and demand. The framework received a lot of attention, as it could explain aspects of the current crisis (as well as the problems in solving it) that were otherwise not understood. A central part of the theory focuses on the process of portfolio decision-making.

In addition to this inner solution within the formalized approach a lot of research in cognitive psychology, anthropology, evolutionary biology, neurology and sociology tries to determine how people actually think (McFadden, 2013). Starting from simplifying assumptions and ideal types has provided invaluable scientific insights (e.g., in game-theory, McCarty and Meirowitz, 2007) but by their innermost definition those theories are very far from "real" thinking. Here, neuroeconomic research has made huge progress in opening up the "black box" of the brain (Camerer et al., 2004).

DellaVigna (2007) provided an overview on the current state of behavioral economics.[420] He focused on deviations from the standard model in terms of preferences, beliefs and decision-making aspects.[421] Following the same principle Walliser (2008) discussed extending the framework of game theory with respect to findings of various laboratory experiments on individual choices and collective interactions.[422] Most attention was paid

[418] Rabin (1998) reviewed various psychological findings on human judgment and behavior. He also pointed out how those can be used to improve assumptions about individual behavior in economic theory.

[419] Ray Dalio is an American businessman and founder of the investment firm Bridgewater Associates (http://www.bwater.com), currently the world's largest hedge fund.

[420] Cf. also Häring and Storbeck (2007) for an entertaining and very informative overview of findings.

[421] Among the aspects considered by DellaVigna (2007) were time preferences (self-control problems), risk preferences (reference dependence), social preferences, overconfidence, projection bias, framing and menu effects, limited attention, persuasion, social pressure and emotions.

[422] Walliser (2008) devoted much space to the fundamental behavioral aspect of bounded rationality which leads to, e.g., financial bubbles, herd behavior or mass hysteria.

to beliefs and cognitive operations of actors in dynamic and strategic environments. The foundations of this "cognitive science" (Walliser, 2008, p. 103) arose from previous work on collective learning processes.

At the moment, a lot of unknown aspects and open questions still exist.[423] A promising approach was put forward by Kahneman (2011), who argued for a combination of two systems of thought: one fast and intuitive, the other rather slow but analytical. The first frequently leads to bad decisions, as people (unconsciously) ignore the second mode because the fast approach is more convenient.[424] While Gladwell (2007) agreed that human behavior is largely marked by mental processes which work rapidly (and automatically), he claimed that people are able to unconsciously filter out the relevant information. This also allows to reach good decisions quickly, even if they are based on relatively little information or experience.[425] Yet, concerning this matter Taleb (2007) prominently discussed the extreme impact of rare events on people's perception and humans' tendency to find simplistic explanations for them in retrospect.

A BROADER RESEARCH FRAMEWORK

One development of this trend was the foundation of the "Institute for new economic thinking" (INET). Its main purpose is to broaden and accelerate the development for replacing current economic theories which were revealed during the recent global financial crisis to be inadequate. Therefore INET promotes research funding, community building, hosting of conferences, etc. The "new thinking" concerns mostly economic topics but is nevertheless important for political science. Looking at its history so far "one could argue that the effect of economics has been felt more strongly in political science than any other social science" (Miller, 1997, p. 1173). It is important for both disciplines to understand how people think and what consequences a procedural change has.

The work of Thaler and Sunstein (2009) constituted a prominent example for the benefit of combining insights from both disciplines. The authors argued in favor of libertarian paternalism, i.e., using behavioral effects to influence choices in benefit of the public good. They based their argumentation on the two systems of thought (Kahneman, 2011) as well as on multiple other fallacies and biases in human decision-making discovered in behavioral economics research. Interpreting these behaviors, they derived policy recommendations for public areas such as healthcare, retirement saving, etc.

Another case of a more inclusive approach to economic research represents the nascent "modern theory of economic order" (von Weizsäcker, 2012, p. 6). This framework incorporates insights of political economics on interactions between political and economic institutions (e.g., North et al., 2009). Taking these and future political equilibria into consid-

[423] Kenning and Plassmann (2005) offered an overview of the current state of neuroeconomic research. They also provided a basic introduction into common concepts and methods.

[424] A good assessment of the theory and suggestions for future research can be found in Evans (2003).

[425] Gigerenzer and Gaissmaier (2011, p. 451) consistently declared that "ignoring part of the information can lead to more accurate judgments than weighting and adding all information".

eration leads to diverging policy recommendations from traditional economic expertise that solely aimed for the removal of market failures, inefficiencies and externalities (cf. Acemoglu and Robinson, 2012). The reason for the disagreement is a diverging view on the causes and consequences of economic development. Following the "iron law of convergence" (this term goes back to Larry Summers, Barro, 1996) and the "modernization hypothesis" (Lipset, 1959), "economic development spurs the introduction and maintenance of higher quality institutions, including well-functioning representative democracy" (Barro, 2012, p. 3). In contrast, new approaches put the "causal effect of income and education on democracy" (Acemoglu et al., 2007, p. 27) into question and point out the critical role of political equilibria. Here, "the political system defines the kind of economic rules of the game and the judicial system you have" (North, 2008, p. 27). Even well-intentioned market interventions which increase the economic performance can "change the political equilibrium in a direction involving greater efficiency losses" (Acemoglu and Robinson, 2013, p. 2) society-wide and over time. Admittedly, the focus on political institutions and the integration of institutional and growth economics is in turn subject to criticism (e.g., Sachs, 2012).[426] This debate is not yet decided but nevertheless emphasizes the potential insights of broadening the social research framework.[427]

Overall, the "new thinking" in economics (and social sciences in general) puts more attention to institutional, historical, and psychological factors (Ruckriegel, 2010). The knowledge of related disciplines is considered (e.g., Colman, 2003), cognitive aspects of behavior are incorporated (e.g., Simon et al., 1992) and inter-dependencies are taken into account (e.g., Hodge and Schwallier, 2006). Most importantly, the constitutive assumptions become more realistic and less idealized by including details which were previously omitted (e.g., Ariely, 2009). This is perfectly in line with my investigation of NSP. Those are an important part of the explanation of human behavior. As political science investigates collective decisions, the tools and methods used should, thus, be based rather on reality than simplicity.

8.4 Chapter summary

This chapter evaluates the question if the required effort for including NSP is justified on the basis of the newly found insights. For this purpose, I focus in Sec. 8.1 on methods and data necessary to implement nonseparability. The assessment is placed in the broader context of using abstraction when setting up analytical models in Sec. 8.2. The concept of abstraction enables the analysis of intricate questions by reducing realities'

[426] Sachs (2012) criticized not the general idea of extending the traditional framework. Rather, he contradicted with Acemoglu and Robinson (2012) sole focus on political institutions and, in particular, the aspects which they did not consider as, e.g., "geopolitics, technological discoveries, and natural resources" (Sachs, 2012, p. 2).

[427] It is important to note that this is not just an abstract debate within scientific circles. The ongoing dispute is also closely covered in daily media (e.g., Braunberger, 2013).

complexity. I do not question the general principle, but criticize its (careless) overuse. This becomes particularly clear when it is contrasted with new trends in social science research paradigms. Here, I describe the so called "new thinking" in behavioral research in Sec. 8.3. This development calls for a more realistic implementation and less restricted or stationary view of the social research environment.

As mentioned earlier in this study, I do not argue that every research project should operationalize nonseparability. Although I criticize the excessive application of the simplifying separability assumption, I will not call for the general (and unaudited) application of nonseparability. It is important to note that, "although this may seem a paradox, all exact science is dominated by the idea of approximation" (Bertrand Russell, as quoted by Auden and Kronenberger, 1966). "Life is complex and so we must simplify our analysis to obtain useful insights. The art of research involves creation of simplifications that provide insights based on evidence and observations" (Hinich, 2008, p. 1000). I will not deny that "if we make our models too complex, we may lose our ability to derive useful predictions for empirical evaluation" (Morton, 1999, p. 280). In addition, Rabin (1998, p. 13) pointed out that tractability and parsimony should be guiding principles when aiming to make research more realistic. Also, I concur with Clarke and Primo (2007) in their effort to modernize the use of models in political science. Models should not only be judged by the accuracy of their deductive predictions, but also for their usefulness "in producing empirical generalizations that may serve as a spur to further modeling efforts" (Clarke and Primo, 2007, p. 741).

My ultimate goal is to lay the foundation for further research on and empirical modeling of NSP. The debate in social science on the relevance of nonseparability is just starting. My contribution is dedicated to further research in two aspects. Firstly, it draws attention to the phenomenon of nonseparability and strengthens the awareness for it. My work proves that research in the fields of survey and referendum design, organizational theory and institutional analysis should consider carefully whether or not NSP are relevant. Both answers (yes or no) can be adequately justified using the arguments discussed in the previous chapters. Yet, it is important to consider both options! I am committed to advocating that the nearly unambiguous standard exclusion of nonseparability in analytical research should end. NSP need not be included into every research project but evaluated with respect to their scientific insights. If they are left out, the reasons for this decision should be stated and consequences for the obtained results discussed.

Secondly, if a research project decides to incorporate NSP, this study comprises a detailed overview of previously conducted theoretical as well as empirical research. I provide empirical tools and guidelines for the operationalization of nonseparability. In addition, I highlight concrete starting points for its assessment; e.g., the reliability of model predictions and changes in actor behavior. My study summarizes various contributions and combines knowledge from different fields. This assists in building a better general insight and specific understanding of NSP.

9 Conclusion and outlook

This chapter concludes my study by summarizing the previous chapters. I outline in Sec. 9.1 the individual steps of my research and highlight my findings. Sec. 9.2 addresses possible future extensions. Those include specific substitutions within my experimental design as well as an alternative focus for research on NSP.

9.1 Final summary

My research addresses the question of whether NSP are of significance for analytical political science. Chap. 1 introduces this research question, demonstrates its relevance and explains the applied research design and method. It also discusses previous contributions to clarify the existing research gap. Whereas the theoretical concept of nonseparability is relatively old, empirical tests are rare. So far, almost all research has used the simplifying assumption of separable preferences as virtually standard. This is most evident when thoroughly considering the amount of previous research on NSP (cf. Sec. 1.2) compared to remaining contributions. I list the scientific fields which gain from research on NSP and describe the further outline of the study. In particular, I point out the expected benefits from the implementation of a laboratory experiment.

My study is devoted to clarifying the relevance of NSP. In Chap. 2 I provide a multitude of examples; all of them strengthen my claim that NSP are commonplace. I use simple hypothetical patterns, genuine empirical examples as well as short case studies to support my argumentation. Next, I particularly evaluate the interplay of institutions and preferences and clarify the importance of nonseparability in this context. Common organizational proceedings as, e.g., delegation, decentralization and specialization are part of the unanswered puzzle of the influence of NSP; this further underlines the relevance of a thorough investigation. Building upon these explanations I describe the theoretical modeling of NSP. With respect to reciprocity I add an additional theoretical aspect to its basic definition. Discussing existing contributions, I put forward arguments about the consequences when neglecting nonseparability in analytical research. Firstly, ignoring NSP leads to a misspecification of actors' utility functions. This implies biased and inaccurate results of models relying on these functions. Also, when decisions affected by NSP are separated, institutional aspects influence decision-making. Thus, secondly, if decisions are taken sequentially, nonseparability favors the actor deciding first. Thirdly, if decisions are taken simultaneously, nonseparability causes sub-optimal outcomes, because strategic voting becomes increasingly difficult.

Next, Chap. 3 is concerned with the misspecification of actors' utility functions. For this purpose I use the field of legislative decision-making in the context of EU politics. More precisely, I undertake empirical analyses to quantify the distortions by examining the impact of NSP on the performance of several legislative decision-making models. These models rely on utility functions which have to be specified according to the policies in question. When a proposal consists of multiple dimensions, the (non)separability of single issues has to be taken into account as otherwise the utility functions are misspecified. The empirical calculations were based on the collected data from the DEU project (Thomson et al., 2006). The coding of this prominent data set for nonseparability demonstrated that the majority of (multi-issue) EU politics are indeed affected by NSP. Moreover, in the majority the respective decisions the effect is non-reciprocal. Thus, the extension of the NSP concept with respect to reciprocity (cf. Sec. 2.4) is strongly encouraged. Next, I investigate if the comparison of various models' predictive accuracy depends on the existence of NSP. Applying simulation techniques, I demonstrate that overlooking NSP may have caused a substantial bias in the empirical evaluation of competing models of EU legislative politics.[428] This confirms my first key question: namely, the importance of a correct model specification with respect to nonseparability to avoid invalid conclusions. In other words, neglecting NSP poses a threat to the inference of the corresponding research. Overall, by comparing my results to previous research I prove the impact and evidence of NSP with real-world data.

The second part of the study (which comprises Chap. 4 to Chap. 7) investigates the implications of NSP for individual and collective decision-making. Here, I rely on a laboratory experiment, the design of which is explained in detail in Chap. 4. I argue that collective decision-making in committees is well-suited for laboratory scrutiny, as the level of monitoring and the focus on processes in decision-making correspond well to both laboratory and real-world politics (Levitt and List, 2006). Also, as I perform a validity test on an existing, theoretically well-known but empirically neglected concept, the experimental method fits my purpose perfectly (Schram, 2005).

Following Levitt and List (2006, 2007) I emphasize the qualitative patterns of my findings. Because of the specific laboratory environment I refrain from singling out the specific parameter estimates (e.g., for social welfare) as my main findings. Chap. 5 looks into the results of my laboratory experiment on the aggregated and Chap. 6 on the individual level. The main findings of the experiment are fourfold. Firstly, the performance of the deterministic equilibrium concept is rather low. The probabilistic extension of the core at the individual level derived in Sec. 6.3.2 is far more appropriate. This indicates clearly that also in collective decisions behavioral aspects such as risk aversion and so-

428 The bias arises from the different degree of restrictions incorporated into the models. This is particularly unsatisfactory for those researchers interested in the effect of procedural aspects.

phistication must be modeled at the individual level. Such a comparison of collective and individual level analysis may allow new insights even into already familiar concepts. An argument in favor of this juxtaposition was the frequent occurrence of minimal-winning coalitions. The information provided by the outvoted subjects is only accessible at the individual level. On the other hand, the collective level might be blurred by the noise of single seditious individuals (cf. Goeree et al., 2002).

Secondly, the long-established trade-off between efficiency (decision costs) and effectiveness (allocation and distribution of wealth) in collective decision-making (cf. Buchanan and Tullock, 1962) is the focus of Chap. 5. I found that, as expected, delegation increased efficiency in terms of decision-making speed. The insights on effectiveness are more ambiguous as delegation provided more stability and fewer imbalances. However, as predicted by theory, with respect to social welfare allocation the pooled decision-making procedure delivered the best performance, whereas unaccounted NSP led to sub-optimal outcomes and Pareto-inferior decisions. This highlights the necessity of well-coordinated delegation. In other words, the organizational structure must enable actors to assess their behavior's impact on the final outcome to prevent less than optimal behavior.

Thirdly, institutions are not neutral, because different decision mechanisms led to substantially different outcomes. This confirms my second key question that nonseparability in conjunction with the respective institutional arrangement indeed influences collective as well as individual behavior. In the specific case of my investigation, unaccounted nonseparability increased sincere behavior; at the same time sophisticated social considerations are replaced by simplistic self-interest. This indicates an interaction among institutions and preferences which calls the neo-classical view in economics (Stigler, 1950) that preferences are exogenously fixed (i.e., that behavior is driven by hedonic utility) into question. This is in line with current cognitive science and psychologists' findings that "people's decisions can be highly sensitive to situational factors, even when such factors are unrelated to the actual utility of that course of action" (Ariely and Norton, 2008, p. 13). The task for the subjects in my experiment was marked with a high level of complexity and uncertainty; these two aspects accompany every decision affected by NSP. Here, my design differs from standard laboratory experiments which typically use simple games and setups.[429] If one does not account for the possible influence of an institutional setting when measuring individual preferences, this bears the risk of mistaking the respondent's conditional response for their genuine preference. Disregarding conditional preferences at the individual stage jeopardizes the conclusions drawn at the collective level.

Fourthly, the first-mover advantage in delegation settings based on standard game theory expectations was not found in my experiment. On the contrary, being part of the

[429] The term "simple" is in fact quite often used by experimenters to describe their setup; e.g., Berl et al. (1976) tested the core concept in a "simple n-person cooperative nonsidepayment game", Charness and Rabin (2002) looked into social preferences "with simple tests" while Ert et al. (2011) used "simple extensive form games" for the same purpose and Chou et al. (2009) emphasized the control over game form recognition when playing "a simple two person guessing game".

first stage when deciding on a nonseparable, but nonetheless split, problem resulted in a below average performance. The high level of complexity and uncertainty prevented the subjects from performing the necessary computations. As both aspects varied across decision procedures, the experimental results clearly showed the crowding out of sophistication. Subjects were not able to strategically comprehend the effects of their votes. The simplifying assumptions that individuals can at any rate or in any situation order their preferences amongst a variety of alternatives and choose the optimum is therefore particularly problematic with regard to NSP. This would imply an artificially high level of computational skills (Sen, 1997).

The voting experiment was accompanied by a post-experiment survey, which is analyzed in Chap. 7; it contained 'yes or no' questions as well as free-input fields. Most importantly, the answers obtained confirm the validity of my design. Describing their own voting behavior, subjects stated that they focused on a limited number of criteria when making their decision. This was mainly, not surprisingly, their own payoff. Looking at sophisticated voting the participants took the probable vote of co-players into account and maximized their conditional utility. This speaks strongly for the introduced probabilistic core concept. The alternation of the decision-making procedure influenced the motives of the participants. Also, the comments prove the desire of subjects to communicate. Collective decision-making is a social exercise, and subjects wondered why communication was excluded for such a task.

Finally, Chap. 8 takes the expense for the implementation of nonseparability into consideration. In particular, I discuss the trade-off between feasibility and accuracy in analytical research. Here, I emphasize the indispensability of abstraction, as it enables the analysis of intricate questions by reducing the complexity of reality; but I also criticize its overuse. With respect to NSP I evaluate the necessary effort for incorporation as well as the knowledge gained. I place my argument to consider nonseparability in the context of a current development in social science research paradigms. This trend calls for a more realistic and less idealized approach when investigating human behavior. This also applies to the nonseparability of preferences, which constitutes an undeniable part of human decision-making.

To summarize, this study comprises a multitude of references, various examples of implementation and a comprehensive theoretical discussion on the concept of nonseparability. I demonstrate that including NSP in analytical research prevents biased and inaccurate results. In particular, studies with respect to institutional structures should be careful because the results obtained are not necessarily due to the organizational aspects, but the denial of nonseparability. I do not conclude that NSP has to be considered in all research, because tractability and parsimony are also guiding principles for a more realistic research paradigm. Yet, if the simplifying assumption of separability is made, the reasons should be given and possible consequences discussed. My work provides the following contributions with the required toolbox.

9.2 Future research

In this study I discussed many important aspects when designing and conducting laboratory experiments. The suitability of this approach to my specific research questions is verified in Sec. 1.3.2 and its general abilities are laid out in Sec. 4.1. Also, Sec. 4.2 explains in detail my design choices and their consequences. The appropriate research method depends on the specific research question (Draper, 2004). However, an absolutely perfect, flawless experiment does not exist. I chose a laboratory experiment due to its high level of internal validity when looking at implications of NSP for individual and collective decision-making. The control capabilities of the laboratory facilitated the operationalization. It also made it possible to identify changes in subjects' behavior when confronted with (neutral) decision problems affected by NSP.

As "the proof of external validity is always empirical" (Morton and Williams, 2010, p. 196) it is necessary to complement results of laboratory research with further studies. This calls for further experiments with variations in design, target population and subject recruiting. Yet, it also emphasizes necessary variations in terms of the experimental method. As always, the knowledge gained from scientific research is confined to a certain extent by the limitations of the applied method. Many missing clues can be attributed to specific approaches. Thus, additional research can provide fruitful new insights.[430]

The following sections discuss three possible aspects of future research. Sec. 9.2.1 picks up the discussion on common problems of laboratory subject pools and evaluates the potential of crowdsourcing to overcome these criticisms. Sec. 9.2.2 focuses on the problem of external validity and argues in favor of supplemental field experiments. Finally, Sec. 9.2.3 offers an alternative research focus on coordination in collective decisions. This list does not mean that many more opportunities for future research do not exist; however, with regard to my own study, these are the most noteworthy and supplemental.

9.2.1 Crowdsourcing

In Sec. 4.2.9 I discuss in detail the shortcomings of laboratory experiments with respect to subject recruiting. These are that

- all participants volunteer for the task.
- nearly all subject pools consist to an overwhelming extent of students.
- the participants might know each other.

[430] An excellent example of how to "prove" external validity can be found in Bosch-Domenech et al. (2002). The authors performed a meta-study of a previous beauty-contest conducted in laboratory as well as in newspaper experiments. In addition, they themselves collected more data in classrooms, conferences, by e-mail, or through newsgroups. This provided a rich variety of different subject pools, sample sizes, payoffs and environmental settings.

Thus, "a random American undergraduate is about 4,000 times more likely than an average human being to be the subject of such a [laboratory] study" (Experimental-Psychology, 2012, p. 69). This really does not account for the principle of random sampling.[431] The problems of self-selection and convenience sampling have been discussed often and to a great extent (e.g., Cleave et al., 2010). Although current research suggests that students are an appropriate subject pool (Branas-Garza et al., 2012), a relatively new approach circumvents these problems and offers interesting new opportunities. Here, the advances in computer technology (e.g., lower costs, faster networks and wider distribution) can also be used for experimental work. A special development is "crowdsourcing"[432], the exploitation of online labor markets. Such a market is an online forum where employers (or researchers) post jobs and employees (or participants) choose which tasks to do for a payoff.[433] It is important to understand that these markets are different from just combining multiple laboratories or subject pools.[434]

Horton et al. (2011) judged online labor markets to be the logical next step of technological improvements in experimental research. Researchers in the early 1990s for the first time had well-developed tools for conducting experiments over local computer networks. The capabilities of the World Wide Web could now be used to free research from the logistical limitations of physical laboratories (cf. Suri and Watts, 2011).

Several firms currently offer their online services for tasks like answering surveys, analyzing texts, ranking websites, etc.[435] The "most popular for scientific purposes is Mechanical Turk, which is run by Amazon" (Experimental-Psychology, 2012, p. 69). Although the markets are still dominated by Americans, they are growing fast world-wide. This implies a much larger and more diversified subject pool which is a clear improvement compared to the "standard" laboratory participant who fits the WEIRD[436] characteristics (Henrich et al., 2010). However, the participants of these platforms "are still a pretty skewed sample of humanity. In particular, they are younger and more liberal than people at large" (Experimental-Psychology, 2012, p. 69). Buhrmester et al. (2011, p. 3) concluded that Amazon Mechanical Turk (AMT) "participants are slightly more demographically diverse than are standard Internet samples and are significantly more diverse

[431] It is also not at all a new phenomenon. Already Reips (2000, p. 92) pointed out that "80% of all psychological studies are conducted with students, while only about 3% of the population are students."

[432] The notion crowdsourcing is used as "an umbrella term for a variety of approaches that harness the potential of large crowds of people by issuing open calls for contribution to particular tasks" (Geiger et al., 2012, p. 2). It is clearly related to the traditional concept of outsourcing, the handover of business functions and structures to third party companies. Yet, for crowdsourcing the voluntary principle is also important (cf. Hammon and Hippner, 2012).

[433] The comparison between laboratory and internet is not new to psychological research (for a comprehensive overview cf. Birnbaum, 2000). Yet, the interest of economics was sparked rather recently in accordance to the rise of online labor markets.

[434] Cf. Frei (2009) for a detailed overview.

[435] These are, e.g., Amazon Mechanical Turk (`https://www.mturk.com`), CrowdFlower (`http://crowdflower.com`), Elance (`http://de.elance.com`), Freelance (`https://www.freelance.de`), Guru (`http://www.guru.com`), oDesk (`https://www.odesk.com`), etc.

[436] The term WEIRD stands for "Western, Educated, Industrialized, Rich and Democratic" and characterizes the land of origin for most laboratory subjects. This is clearly not globally representative.

than typical American college samples." Thus, despite being more diverse than student samples, the markets do not solve all recruiting problems.[437]

All experimental laboratories are equipped with a subject-recruiting mechanism.[438] Most of the time, this is just a database containing a huge number of registered subjects including demographic and contact information. The success of a laboratory (and of the experiments conducted in it) almost always depends on a well-maintained, up to date and, in particular, accurate database. The same is true for online experiments. But without personal contact it is more difficult to ensure reliability. Online labor markets rely on reputation systems (Resnick et al., 2000). Market operators also conduct bank account checks (Horton et al., 2011, p. 6). Overall, the problems of online laboratories are the same as those recruiting databases have. It is difficult and demanding to keep them accurate.

Rand (2012) argued strongly in favor of mechanical turks. By reviewing replication studies on AMT of previous laboratory contributions, he verified that the method is valid: the results revealed the same effects under both conditions. Horton et al. (2011) reached a similar conclusion with respect to internal and external validity. The author also conducted truth checks of self-reported demographics. Depending on the variable the level of reliability was between 81% and 98%.[439]

An argument often made in favor of the online markets is that they are relatively cheap. Mason and Suri (2012) offered practical advice for conducting experiments in this way. Such contributions are much-needed as current reports express concerns that participants are exploited due to small salaries of less than $2 per hour (Experimental-Psychology, 2012, p. 70). Paolacci (2012) operates a blog that has similar goals. While promoting the approach for experimental research in general, his objective is to establish basic guidelines for freelance work.

Overall, conducting experiments in online labor markets is still an unusual approach. However, it offers easy access to a large and more diverse subject pool. As it is also a low cost approach, it increases the possible number of observations. This provides experimental research in political science with a promising opportunity because real-world politics usually affects, or is created by, many people (e.g., Miller, 1997; Mueller, 2003). In general, well-designed laboratory experiments may be limited to a small number of

[437] Reips (2000, p. 96) expected the "Demographics of Internet users [...] to rapidly approach similarity with demographics of the general population." Although this argument was made some time ago and has not yet come completely true it might for future research change the diversity argument significantly.

[438] The Laboratoire d'Economie Experimentale de Montpellier offers a list of experimental laboratories around the world. It can be obtained at `http://leem.lameta.univ-montp1.fr/index.php?page=liste_labos&lang=eng`.

[439] Typical findings of laboratory experiments which have been replicated using online labor markets were, e.g., priming and framing effects (Horton et al., 2011). In general, many current studies look for behavioral differences between experiments conducted in the laboratory and via the internet. Bosch-Domenech et al. (2002, p. 1687) referred to this as "to test the critical assumption of 'parallelism' between the lab and the field." Those contributions included e.g., trust games (Fiedler and Haruvy, 2009), ultimatum games (Goerg et al., 2007), beauty-contests (Selten and Nagel, 1998) and auctions (Lucking-Reiley, 1999). Sec. A.17 contains a more extensive list of references on studies with experiments conducted online and in the laboratory collected by Israel Waichman (Heidelberg University).

observations due to the high internal validity of the research method (cf. Sec. 1.3.2). But political science asks specific questions about collective processes and structures. A good example is the realm of public choice, where "the use of laboratory experiments [...] has also increased rapidly in the last thirty years" (Schram, 2002, p. 1). Many of these experiments have examined various variables and their influence on the provision of public goods (for a detailed overview cf. Ledyard, 1995). While in reality public goods are generally provided within a large population (e.g., pension system and national defense), few experiments investigated the effect of group size. While Isaac et al. (1994) found that public good contributions increase with group size, Offerman et al. (1996) concluded that voluntary contributions decrease with more group members.[440] More important than these controversial insights is the fact that, the usual number of participants in public good experiments was only four to five, and very few studies looked at groups of more than ten (Schram, 2002, p. 2). This shortcoming can be attributed to a large extent to the spatial and organizational limitations of standard laboratories. In addition to the obvious problem of equipping a lab with 200 or more seats,[441] a large number of participants require exponentially greater financial resources. These are exactly the types of logistical problems of standard laboratory experiments in which crowdsourcing might help.

It is also important to note that subjects do not necessarily realize that they are part of a research project. They are not aware that their "employer" might be a researcher. As the interaction takes place in their usual (online) working environment this can prevent a Hawthorne effect[442], where subjects change their behavior because they are aware of being monitored. This is even truer if the use of online labor markets follows the example of the conversion of people's social life into online networks (e.g., Facebook).

The potential of linking theoretical work to a broad empirical analysis is not in question. However, online experiments have limitations and pitfalls of their own (Eckel and Wilson, 2006). It is much harder to conduct synchronized experiments and to secure high quality work. When you only receive coded input online, a subject's identity remains anonymous (raising the question who actually sat in front of the screen). In particular, communication between subjects can become a critical issue. The laboratory offers a much higher degree of control (Morton and Williams, 2010, p. 31ff). In addition, typical online data problems have to be taken into account, such as how to keep data private, how to maintain code security, etc. Mason and Suri (2012, p. 11ff) discussed these problems in detail and suggested possible improvements. However, it looks as though the solution simply exchanges the shortcomings of one approach for the shortcomings of an-

[440] More specifically, Isaac et al. (1994, p. 32) pointed out that not group size per se matters, but the interaction between group size and the marginal per capita return from the public good (i.e., the marginal rate of substitution).

[441] To give an exemplary insight on the possible number of participants, the study of Diederich and Goeschl (2011) investigated the determinants for charitable giving and is based on a large-scale field experiment with 2,440 subjects. To include this number of persons successively, let alone simultaneously, into a laboratory experiment is infeasible.

[442] Gillespie (1991) provided an excellent review of the Hawthorne experiments.

other. Tab. 9.1 summarizes the arguments discussed so far, both in favor of and against internet experiments.

Table 9.1: Pros and cons of internet experiments

EXPLANATORY NOTE
The table lists the pros and cons of conducting experiments online instead of in the laboratory. It also cites exemplary sources as evidence for the arguments made. The illustration is based on Charness et al. (2007a, p. 91, table 1).

Arguments against internet experiments	Arguments in favor of internet experiments
• Not everyone has internet access (Krantz et al., 1997)	• More diverse populations (generalizability) (Buhrmester et al., 2011)
• More noise and higher variance (Shavit et al., 2001)	• Demographics approach those of the general population (Kehoe and Pitkow, 1996)
• Subjects appear less attentive in Internet experiments (Anderhub et al., 2001)	• Reduced experiment(er) effects as the web might rather reflect a "natural environment" (Harrison and List, 2004)
• Subjects fear deception (e.g., subjects do not believe they are matched with a real person) (Eckel and Wilson, 2006)	• Lower costs and fewer physical limitations (Suri and Watts, 2011)
• Loss of control over the physical environment (Morton and Williams, 2010)	• No discernible differences in levels of rationality (Nagel et al., 2002)

The main reason to discuss crowdsourcing in the context of NSP is its ability to facilitate cross-cultural research (Paolacci et al., 2010).[443] Various contributions (experimental as well as non-experimental) suggested that preferences and attitudes are different across cultural backgrounds (e.g., Fiorino and Ricciuti, 2007; Bohnet et al., 2008). Using a survey designed accordingly, Lacy (2001b) identified in the U.S. numerous NSP between political issues: the preferences on income taxation depend on crime prevention policies, preferences on environmental pollution depend on environmental regulation, preferences on defense spending depend on social spending, preferences on immigration policy depend on the constitutional status of English being the only official language, etc. But do the same dependencies exist in other countries as well? More likely than not the political agenda will consist of other topics (Ehmke et al., 2005; Goerg et al., 2007). Those issues might be separable, but as this study has shown, this should not be taken for granted. To provide a list with nonseparable policy fields across countries would be a good way to increase the awareness of NSP. It would also offer valuable insights for research on elections and campaigns. Of course, for such statements a representative sample is inevitable; whether crowdsourcing can provide this to a sufficient degree, however, stays questionable but seems worth investigating.

9.2.2 Going to the field

The idea of bringing research on NSP "to the field" (Harrison and List, 2004) follows similar argumentation. Compared to the laboratory setting, a field experiment inhibits

[443] Cf. Eriksson and Simpson (2010) for an example of such a study. The authors conducted an online-survey testing attitudes of risk preferences in the U.S. and India.

a lower degree of environmental control. Many more possibilities and distorting influences on subjects have to be considered. Nevertheless, Levitt and List (2007) pointed out that lab-generated as well as observational data suffer from shortcomings. Here, a "well-designed field experiment [...] can serve as a bridge connecting these two empirical approaches" (Levitt and List, 2007, p. 171). If this is done correctly, a researcher may be able to observe a subject's natural behavior.

Studying a subject outside the "sterile environment" (Harrison and List, 2004, p. 1009) of a laboratory has proven to be valuable many times.[444] Such contributions complemented existing laboratory investigations in various fields very well. Bahry and Wilson (2006) evaluated fairness in ultimatum games, Carpenter and Seki (2005) searched for social preferences, Holm and Nystedt (2010) investigated collective trust behavior, Olken (2008) analyzed the provision of public goods, Cohen and Nisbett (1997) looked for reasons for the perpetuation of honor-related violence, etc.[445] Going more into specific topics of politics, Wantchekon (2003) assessed clientelism, Kuklinski et al. (1997) examined the impact of racial prejudice, Gerber et al. (2009) scrutinized media influence on voting, etc. I list these contributions in such detail for two reasons. Firstly, they reflect a variety of research questions which profited from field investigations. There is no argument why this should not also be true for research on NSP. Secondly, some of them (e.g., racial prejudice) may very well be affected by NSP themselves.

It is important to note that the term "field experiment" does not just refer to a physical location, and by no means are the methodological requirements lower or less stringent. With the focus on "lab in the field" experiments Hamenstädt (2012) highlighted the potential problems of self-selection and non-compliance and emphasized the importance of randomization. Harrison and List (2004) broadened the definition of field experiments considerably. They explained that it involves the nature of the subject pool, the information subjects bring to the task and the type of commodity and stakes as well. In their taxonomy a natural field experiment takes "place in the subjects' natural environment and, importantly, where the subjects do not know they are in an experiment" (Morton and Williams, 2010, p. 223). Depending how far one views online labor markets as being a natural environment, crowdsourcing (cf. Sec. 9.2.1) might belong into this category.

9.2.3 Communication and interaction

The laboratory empowers the researcher with a high degree of environmental control. Nearly every aspect of subjects' interaction and perception can be modified (cf. Morton and Williams, 2010, Chap. 4). Communication is a central aspect of human interplay in

[444] In political science quasi-experiments are also common (Bernauer et al., 2009, p. 91). In this section, it is not my intention to distinguish between experiments and quasi-experiments; rather I will focus on the differences for external validity when leaving the laboratory. For a comprehensive overview on experimental and quasi-experimental designs cf. (Shadish et al., 2002).

[445] Cf. Ortmann (2005) for an overview on field experiments in Economics and some methodological notes.

general and many studies have proven its influence on behavior and beliefs (e.g., Andreoni and Rao, 2011). It is also a fundamental aspect of group coordination. Current research in public administration theory, organization theory and planning theory turns more and more towards its dynamic aspects (cf. Pedersen et al., 2011). Diverging from traditional theories, these new contributions are "pointing to the relational, interpretive, interdependent, and interactive aspects of all coordination processes" (Pedersen et al., 2011, p. 375).

In previous laboratory experiments two ways of dealing with communication could be observed. One approach was to control communication "in order to prevent possible cross-effects that occur with communication. [..] In almost all game theory experiments, communication between subjects is not allowed except under particularly controlled circumstances" (Morton and Williams, 2010, p. 122). This cautious handling was due to the less predictable confounding of communication with other variables of the experiment.

A second stream of research used communication as its central treatment variable. For example, Cason and Mui (2007) found out that different forms of (non-binding) communication facilitated collective coordination. Isaac and Walker (1988b) looked for similar effects when analyzing group size effects in public good experiments. Miller and Oppenheimer (1982) found that communication enforced the relevance of universalism. Changes in behavior were also observed by Muren and Pyddoke (1999) after their subjects were allowed to interact and build up trust in each other. In summary, these studies agree that social interaction enhances social consideration.[446] The literature is strongly linked with work on framing effects (Tversky and Kahneman, 1981). Morton and Williams (2010, p. 28) stated that "framing effects work when a communication causes an individual to alter the weight he or she places on a consideration in evaluating an issue or an event."[447]

Of course, interactions did not exclusively lead to cooperative aspects. Eckel and Petrie (2011) found that information is used to discriminate between people. They revealed that a specific "face, it appears, has strategic value" (Eckel and Petrie, 2011, p. 1497). In his experiment Vincent (2012) informed his subjects about the name of their experimental partners, hereby disclosing gender and race. His results showed no general trend but suggested that "racial group behavior varies depending on the racial composition and attitudes of the population being observed" (Vincent, 2012, p. 1).

While these examples disagreed on intensity and shape, all studies agreed that interaction influences behavior. Sunitiyoso et al. (2011) investigated the net and gross effects of social interaction as well as the direction of change of individual behavior. They identified behavior which conforms to models of social learning. Mickan and Rodger (2000,

[446] However, as discussed in Sec. 4.2.2 and Sec. 6.1.2 social motivations may be enhanced by communication, but they do not dependent on it (Fehr et al., 2002a).

[447] Druckman (2004) identified various criteria for the success of framing. He concluded that a multitude of contextual forces as well as individual attributes are important.

p. 205) argued that communication is one of the most important factors for group efficiency. Allen and Fusfeld (1974) found that communication networks as well as seating arrangements influence the frequency and patterns of interaction. They concluded that "communication is influenced by the physical, architectural arrangement of the laboratory. Communication between individuals is very sensitive to both the horizontal and vertical distances separating them" (Allen and Fusfeld, 1974, p. 39). But what happens if subjects are given a choice? Lai and Lim (2012) carried out a test to uncover if subjects communicate when given the (costly) possibility to do so. They let subjects chose between delegating competence or communicating with other subjects. The authors observed "significantly more choices of delegation than of communication" (Lai and Lim, 2012, p. 541). However, subjects still under-delegated as they hoped to profit from information of their co-players.

I implemented my laboratory experiment with the z-Tree software. This program also includes a chat feature which enables subjects to send text messages to each other (cf. Fischbacher, 2007, p. 172). Most important, the experimenter would still control and observe the content of the messages, the degree and frequency of contact, etc. So, in general, monitoring communication between subjects would be easy to achieve. A much more demanding task is to develop a corresponding theory and to translate it into a specific protocol. Communication in collective decision-making means more than merely agreeing on one alternative. Committee meetings possess multiple levels of mutual influence. Communication involves an observable interchange of information and subtle interactions of power, attitudes and values (Loxley, 1997). Therefore, it is difficult to decide which kind of communication should be allowed.

One aspect refers to the technical implementation, e.g., should the discussion be based on text alone or should it also be visual; should the contact take place in person or via a screen? In their collusion games Cooper and Kühn (2009) observed clear differences when allowing either unlimited pre-game communication or limited message-space during the experiment. Dutcher (2012) investigated the effects of telecommuting on productivity and found that it has positive implications with creative and negative implications with dull tasks. Thus, the kind of experimental task may also interact with communication and the different technical implementations. Another aspect concerns the structure of communication. Should everybody be able to talk to everybody else or should one host structure the discussion?[448] Previous public good experiments investigated the influence of a group leader on contributions. They found that one-sided communication increased the contributions although a lot of cheap-talk was generated (e.g., Güth et al., 2004; Koukoumelis et al., 2012). It would be interesting to observe how this kind of communication would affect my experimental setting and its distributional patterns. Yet, an unanswered matter in public good experiments is the effect of how the group leader is

[448] Kriss and Weber (2012) provided a survey of laboratory insights on the formation of groups with(out) leaders and the shaping of economic organization.

determined. The leadership role can be assigned exogenously by the experimenter or endogenously by the participants[449] (cf. Arbak and Villeval, 2007).

Overall, thinking about communication is more than just recording words; in the laboratory it is possible to do more. When I mention communication as a possible part of future research, I do not simply refer to a degree of information exchange but focus on if and how the subjects use their interpersonal contact to influence the collective decision. The pending task is to scrutinize how separating the decision over nonseparable issues affects communication (and behavior).[450]

Weihe et al. (2008) looked at the principles of committee deliberation processes.[451] Deliberation means to scrutinize the complete process of negotiations and to focus on the establishment of liabilities in committees (Gutmann and Thompson, 2004). This is contrary to the traditional rational choice approach (starting with the contributions of Black, 1958, 1991) which does not look into the concrete course of (internal) committee meetings. Thus, the focus on liability is in fact a theoretical principal and follows the perception of politics, which are seen as emerging from binding decisions (cf. Easton, 1957; Scharpf, 2000).

To evaluate the deliberative process it is necessary to provide an extensive documentation of the communication process and the various interaction steps (Weihe et al., 2008, p. 344). Such an inquiry will be difficult and laborious. Weihe et al. (2008) presented a three-piece theoretical scheme for covering the process. They dispersed the unification process into the steps of proposal, acceptance and confirmation. The course of the discussion had then to be sorted into these categories enabling the analysis. Of course, this did not yet deal with the concrete operationalization of the measurement.

Again, here the high degree of environmental control of the laboratory can help. Hennig-Schmidt (1997) used a video analysis to investigate the break-offs in bargaining negotiations. In a similar fashion Nullmeier and Pritzlaff (2009) analyzed the dynamics and power relations of committee decision-making. The difficult part is to set up an appropriate coding scheme to transcribe the video material into a specific score notation[452]. This notation must then be divided into a small section and thoroughly scrutinized to enable a process analysis of the deliberative action (Weihe et al., 2008, p. 344ff).

Looking into group decision-making at this level of detail enables researchers to discover insights into the behavior of individuals. How do they react when the settings change? Which attitude do they adopt when talking? If distinguishing between sincere and sophisticated behavior, which attitude do they chose? We know that the terms we use, and how we use them, influences our perception and our actions (Pritzlaff, 2006). But how

[449] This can be done in various ways: through elections, based on the performance in solving a puzzle or answering knowledge questions, etc.

[450] Cf. Bardsley et al. (2010) for a more general assessment of the current state and future development of experimental (economic) research.

[451] For a comprehensive overview on deliberative democracy cf. Bohman and Rehg (1997) and Elster (1998).

[452] Cf. Weihe et al. (2008, p. 347) for an example of such a score notation.

do individuals form and shape collective decisions? Furthermore, in the context of NSP, how does separating the decision over nonseparable issues alter the findings.

Using video analysis is only one possible way to monitor communication and interaction in an experiment. There are many more, and often chat protocols or pictures might be perfectly adequate. Nevertheless, group decision-making blurs the ability to infer from the observed behavior. To cover as many different levels of human interaction as possible might be a laborious but appropriate reaction to obtain validity.

Bibliography

Abbink, K. and H. Hennig-Schmidt. 2006. "Neutral versus loaded instructions in a bribery experiment." *Experimental Economics* 9:103–121.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2007. "Reevaluating the Modernization Hypothesis." *NBER Working Paper* 13334.

Acemoglu, Daron and James A Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business.

Acemoglu, Daron and James A Robinson. 2013. "Economics Versus Politics: Pitfalls of Policy Advice." *NBER Working Paper, SSRN* w18921.

Achen, C. 1983. "Towards Theories of Data." In *Political Science: The State of the Discipline*, edited by A. Finifter, chapter 3, pp. 69–93. Washington, DC: American Political Science Association.

Achen, Christopher H. 2006. "Evaluating Political Decision Making Models." In *The European Union Decides*, edited by R. Thomson, F. N. Stokman, C. H. Achen, and T. König, pp. 264–298. Cambridge: Cambridge University Press.

Ackert, Lucy F., Ann B. Gillette, Jorge Martinez-Vazquez, and Mark Rider. 2009. "Risk Tolerance, Self-Interest, and Social Preferences." *mimeo. Experimental Economics Center Working Paper Series, Georgia State University.* http://excen.gsu.edu/workingpapers/GSU_EXCEN_WP_2009-04.pdf (accessed March 2, 2012).

Akaike, H. 1973. "Information theory and an extension of the maximum likelihood principle." In *Second International Symposium on Information Theory*, edited by B. N. Petrov and B. F. Csaki, pp. 267–281. Academiai Kiado: Budapest.

Akerlof, G. A. 2007. "The missing Motivation in Macroeconomics." In *Presidential Adress at the 199th Conference of the American Economics Association*.

Akerlof, George A. and Robert J. Shiller. 2009. *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press.

Al-Ubaydli, Omar, Uri Gneezy, Min Sok Lee, and John A. List. 2010. "Towards an understanding of the relative strengths of positive and negative reciprocity." *Judgment and Decision Making, Society for Judgment and Decision Making* 5:524–539.

Al-Ubaydli, Omar and Min Sok Lee. 2009. "An experimental study of asymmetric reciprocity." *Journal of Economic Behavior and Organization* 72:738–749.

Al-Ubaydli, Omar and Min Sok Lee. 2012. "Do you reward and punish in the way you think others expect you to?" *The Journal of Socio-Economics* 41:336–343.

Aldrich, Howard E. 1979. *Organizations and Environments*. Prentice-Hall, New York.

Allen, Thomas J. and Alan R. Fusfeld. 1974. "Research Laboratory Architecture and the Structuring of Ccmnunications." *Working papers, Massachusetts Institute of Technology (MIT)*. 692-74.

Ambrose, M. L. and M. Schminke. 2003. "Organization structure as a moderator of the relationship between procedural justice, interactional justice, perceived organizational support, and supervisory trust." *Journal of Applied Psychology* 88:295–305.

Amir, O., D. G. Rand, and Y. K. Gal. 2012. "Economic Games on the Internet. The Effect of 1-Dollar Stakes." *PLOS one* 7.

Amsterdam. 1997. *Vertrag zur Gruendung der Europaeischen Gemeinschaft*. Europaeischer Rat. signed on 2nd October.

Anderhub, V., R. Mueller, and C. Schmidt. 2001. "Design and evaluation of an economic experiment via the Internet." *Journal of Economic Behavior and Organization* 46:227–247.

Anderson, Lisa R. and Charles A. Holt. 2002. "Experimental economics and public choice." *mimeo. Prepared for the Encyclopedia of Public Choice.* http://web.cenet.org.cn/upfile/44224.pdf (accessed July 24, 2012).

Andersson, Ola, Cedric Argenton, and Jörgen W. Weibull. 2012. "Robustness to Strategic Uncertainty." *Working Paper Series , Research Institute of Industrial Economics* 910.

Andreoni, James, William Harbaugh, and Lise Vesterlund. 2003. "The Carrot or the Stick: Rewards, Punishments, and Cooperation." *American Economic Review* 93:893–902. American Economic Association.

Andreoni, J. and J. Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70:737–753.

Andreoni, James and Ragan Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fund-raising." *Journal of Public Economics* 88:1605–1623.

Andreoni, James and Justin M. Rao. 2011. "The power of asking: How communication affects selfishness, empathy, and altruism." *Journal of Public Economics* 95:513–520.

Anonymous. 2010. "Nonseparable Preferences in Public Surveys." *mimeo. Political Studies.*

Arbak, E. and M.-C. Villeval. 2007. "Endogenous Leadership: Selection and Influence." *IZA Discussion Paper* 2732.

Arbuthnott, J. 1710. "An argument for divine providence, taken from the constant regularity observed in the births of both sexes." *Philosophical Transaction of the Royal Society of London* 27:186–190.

Ariely, D. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. Harper Collins.

Ariely, Dan. 2009. "The End of Rational Economics." *Harvard Business Review* .

Ariely, Dan and Michael I. Norton. 2008. "How actions create - not just reveal - preferences." *TRENDS in Cognitive Sciences* 12:13–16.

Arregui, J., F. Stokman, and R. Thomson. 2006. "Compromise, Exchange and Challenge in the European Union." In *The European Union Decides*, edited by Stokman F. Achen C. Thomson, R. and T. König, chapter 5. Cambridge University Press.

Arrow, Kenneth J. 1950. "A Difficulty in the Concept of Social Welfare." *Journal of Political Economy, University of Chicago Press* 58:328–346.

Arzieli, Yaron, Christopher P. Chambers, and Paul J. Healy. 2012. "Eliciting multiple choices: A theoretical analysis of incentives in experiments." *mimeo. Ohio State University.* http://healy.econ.ohio-state.edu/papers/Azrieli_Cambers_Healy-CombinedDecisions.pdf (accessed October 5, 2012).

Auden, Wystan Hugh and Louis Kronenberger. 1966. *The Viking Book of Aphorisms: A Personal Selection*. Dorset Press.

AWI. 2013. "Teilnahme-Regeln (AGBs)." Alfred-Weber-Institut für Wirtschaftswissenschaften der Ruprecht-Karls-Universität Heidelberg, http://www.uni-heidelberg.de/md/awi/professuren/beheco/teilnahmeregeln_awi_lab.pdf (accessed November 20, 2013).

Azaiez, Ines. 2010. "Sampling of Alternatives for Logit Mixture Models." Master Thesis, Massachusetts Institute of Technology, http://transp-or.epfl.ch/documents/masterTheses/AZAIEZ10.pdf (accessed November 28, 2013).

Azar, Ofer H. 2010. "Does relative thinking exist in mixed compensation schemes?" *mimeo. Ben-Gurion University of the Negev.* http://thema.u-cergy.fr/IMG/pdf/AEE_24_06_10.pdf (accessed February 17, 2012).

Bahry, D. L. and R. K. Wilson. 2006. "Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method." *Journal of Economic Behavior and Organization* 60:37–54.

Bailer, Stefanie. 2006. "The European Commission and Its Legislative Activity - Not as Integrationist and Autonomous as Believed." *Working paper of the Center for Comparative and International Studies (ETH Zurich and University of Zurich)* 24.

Bandiera, O., I. Barankay, and I. Rasul. 2005. "Social preferences and the response to incentives: Evidence from personnel data." *Quarterly Journal of Economics* 120:917–962.

Banuri, Sheheryar, Catherine Eckel, and Rick Wilson. 2011. "Deconstructing Nepotism." *mimeo. Working paper, University of Texas, Dallas.* http://cbees.utdallas.edu/papers/DeconstructingNepotism.pdf (accessed September 2, 2012).

Barbera, S., J. Masso, and A. Neme. 2005. "Voting by committees under constraints." *Journal of Economic Theory* 122:185–205.

Bardsley, N., R. Cubitt, G. Loomes, P. Moffatt, C. Starmer, and R. Sugden. 2010. *Experimental Economics: Rethinking the Rules*. Princeton: Princeton University Press.

Barmettler, Franziska, Ernst Fehr, and Christian Zehnder. 2011. "Big Experimenter Is Watching You! Anonymity and Prosocial Behavior in the Laboratory." *Institute for the Study of Labor (IZA) Discussion Papers* 5925.

Baron, D. P. and J. A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83:1181–1206.

Barrett, Richard and Maurice Salles. 2006. "Social Choice With Fuzzy Preferences." *Economics Working Paper Archive, Center for Research in Economics and Management (CREM), University of Rennes* 200615.

Barro, R. 1996. *Getting It Right: Markets and Choices in a Free Society*. Cambridge, MA, MIT Press.

Barro, Robert J. 2012. "Convergence and Modernization Revisited." *NBER Working Paper* 18295.

Basedow, Juergen, Joern Aldag, Justus Haucap, Peter-Michael Preusker, and Katharina M. Trebitsch. 2006. "Die Privatisierung der Deutschen Bahn AG." *Sondergutachten der Monopolkommission gemaess § 44 Abs. 1 Satz 4 GWB* .

Baumeister, R., E. Bratslavsky, C. Finkenauer, and K. Vohs. 2001. "Bad is stronger than good." *Review of General Psychology* 5:323–370.

Bazerman, Max H. and Roy J. Lewicki. 1985. "Contemporary Research Directions in the Study of Negotiations in Organizations: A Selective Overview." *Journal of Occupational Behaviour* 6:1–17.

Beck, Kurt. 2011. "Weniger Soldaten bedeutet auch weniger Kasernen - Kurt Beck ueber die Schliessung von Kasernen und die Zukunft der Bundeswehr." *Deutschalndfunk* 26.10.

Beck, Nathaniel L. 2000. "Political Methodology: A Welcoming Discipline." *Journal of the American Statistical Association* 95:651–654.

Becker, G. 1971. *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.

Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago University Press.

Becker, R. and S. Chakrabarti. 1995. "The Recursive Core." *Econometrica* 63:401–423.

Beckman, Steven R., John P. Formby, James W. Smith, and Buhong Zheng. 2002. "Envy, malice and Pareto efficiency: An experimental examination." *Social Choice and Welfare* 19:349–367.

Beedham, Brian. 1993. "A Better Way to Vote." *150 Economist Years* pp. 5–8.

Beitz, Charles R. 1990. *Political Equality: An Essay in Democratic Theory*. Princeton University Press.

Bellemare, Charles and Sabine Kroger. 2007. "On representative social capital." *European Economic Review, Elsevier* 51:183–202.

Bellu, Lorenzo Giovanni and Paolo Liberati. 2006. "Inequality Analysis - The Gini Index." In *United Nations - Inequality Analysis*, chapter Module 040. Food and Agriculture Organization of the United Nations.

Belot, Michele and Jeroen Van De Ven. 2009. "Friendships and Favoritism on the Schoolground." *mimeo. Nuffield College, Oxford University.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.182.3724&rep=rep1&type=pdf (accessed July 4, 2012).

Benartzi, S. and R. H. Thaler. 1995. "Myopic loss aversion and the equity premium puzzle." *The Quarterly Journal of Economics* 110:73–92.

Bendor, Jonathan, Ami Glazer, and Thomas H. Hammond. 2000. "Theories of Delegation." *Annual Review of Political Science* 4:235–269.

Bendor, J. and A. Meirowitz. 2004. "Spatial models of delegation." *American Political Science Review* 98:293–310.

Benoit, Jean-Pierre and Lewis A. Kornhauser. 1991. "Voting Simply in the Election of Assemblies." *Working Papers, C.V. Starr Center for Applied Economics, New York University.* 91-32.

Benoit, Jean-Pierre and Lewis A. Kornhauser. 2007. *Only a Dictatorship is Efficient or Neutral*, volume 07-01. New York: NYU Law School.

Benoit, K. and M. Laver. 2007a. "Benchmarks for text analysis: A response to Budge and Pennings." *Electoral Studies* 26:130–135.

Benoit, K. and M. Laver. 2007b. "Estimating party policy positions: Comparing expert surveys and hand-coded content analysis." *Electoral Studies* 26:90–107.

Benoit, Kenneth and Michael Laver. 2007c. *Party policy in modern democracies*. Milton Park, Abingdon, Oxon ; New York: Routledge.

Benoit, K., M. Laver, C. Arnold, P. Pennings, and M. O. Hosli. 2005. "Measuring national delegate positions at the convention on the future of Europe using computerized word scoring." *European Union Politics* 6:291–313.

Benz, Arthur. 2005. "Kein Ausweg aus der Politikverflechtung? Warum die Bundesstaatskommission scheiterte, aber nicht scheitern musste." *Politische Vierteljahresschrift (PVS)* 46:204–214.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 2005. "Risk Preference Instability Across Institutions: A Dilemma." *Proceedings of the National Accademy of Sciences* 102:4209–4214.

Bergstrom, Ted. 2011. "Lecture Notes on Separable Preferences." *mimeo. USCB Econ 210A.* http://www.econ.ucsb.edu/~tedb/Courses/GraduateTheoryUCSB/separability notes.pdf (accessed September 22, 2012).

Bergstrom, T.C. and J.H. Miller. 2000. *Experiments with Economic Principles: Microeconomics*. New York: McGraw-Hill, 2nd edition.

Berl, Janet, Richard D. McKelvey, Peter C. Ordeshook, and Mark Winer. 1976. "An Experimental Test of the Core in a Simple N-Person Cooperative Nonsidepayment Game." *Journal of Conflict Resolution* 20:453–479.

Bernauer, Thomas, Detlef Jahn, Patrick Kuhn, and Stefanie Walter. 2009. *Einführung in die Politikwissenschaft*. Baden-Baden: Nomos, 2nd edition.

Bernheim, B. D., B. Peleg, and M. D. Whinston. 1987a. "Coalition-proof Nash equilibria. I. Concepts." *Journal of Economic Theory* 42:1–12.

Bernheim, B. D. and M. D. Whinston. 1987b. "Coalition-proof Nash equilibria. II. Applications." *Journal of Economic Theory* 42:13–29.

Bernstein, Fernando and Awi Federgruen. 2005. "Decentralized Supply Chains with Competing Retailers under Demand Uncertainty." *Management Science* 51:18–29.

Bianco, William T., Ivan Jeliazkov, and Itai Sened. 2004. "The Uncovered Set and the Limits of Legislative Action." *Political Analysis* 12:256–276.

Bianco, William T., Michael S. Lynch, Gary J. Miller, and Itai Sened. 2008. "The Constrained Instability of Majority Rule: Experiments on the Robustness of the Uncovered Set." *Political Analysis* 16:115–137.

Birnbaum, Michael H. 2000. *Psychological Experiments on the Internet*. Academic Press.

Black, Duncan. 1948. "On the Rationale of Group Decision-Making." *The Journal of Political Economy, University of Chicago Press* 56(1):23–34.

Black, Duncan. 1958. *The theory of committees and elections*. Cambridge Eng.: University Press.

Black, Duncan. 1991. "Arrows Work and the Normative Theory of Committees." *Journalof Theoretical Politics* 3:259–276.

Bloomfield, Robert and Alyssa Anderson. 2010. "Experimental finance." *Johnson School Research Paper Series* 23.

Bloomfield, R.J., W.B. Tayler, and F. Zhou. 2009. "Momentum, Reversal, and Uninformed Traders in Laboratory Markets." *The Journal of Finance* 64:2535–2558.

Blundell, R. 1999. "Consumer Demand and Intertemporal Allocations: Engel, Slutsky and Frisch." In *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, edited by S. Strom, chapter 5, pp. 147–166. Cambridge University Press.

BMVG, Bundesministerium der Verteidigung. 2012. "Infrastrukturmassnahmen /-investitionen der Bundeswehr zur Umsetzung des neuen Stationierungskonzepts." *Stellungnahme zur schriftliche Frage beim Bundeskanzleramt am 6. Maerz* 3/50.

Bogart, Dan. 2009. "Nationalizations and the Development of Transport Systems: Cross-Country Evidence from Railroad Networks: 1860-1912." *Journal of Economic History* 69:202–237.

Bogart, Dan. 2010. "A Global Perspective on Railway Inefficiency and the Rise of State Ownership, 1880-1912." *Explorations in Economic History* 47:158–178.

Bohman, James and William Rehg. 1997. *Deliberative Democracy. Essays on Reason and Politics.* Cambridge/London: MIT Press.

Bohnet, Iris and Bruno S. Frey. 1999. "The sound of silence in prisoner's dilemma and dictator games." *Journal of Economic Behavior and Organization* 38:43–57.

Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser. 2008. "Betrayal aversion: evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States." *American Economic Review* 98:294–310.

Bolton, G. E., J. Brandts, and A. Ockenfels. 2005. "Fair procedures: Evidence from games involving lotteries." *Economic Journal* 115:1054–1076.

Bolton, G. E. and A. Ockenfels. 2000. "ERC: A theory of equity, reciprocity, and competition." *American Economic Review* 90:166–193.

Bolton, Gary E. and Axel Ockenfels. 2005. "The Influence of Information Externalities on the Value of Reputation Building - An Experiment." *mimeo. Working Paper Series in Economics, University of Cologne, no. 17.* http://ockenfels.uni-koeln.de/fileadmin/wiso_fak/stawi-ockenfels/pdf/wp_series_download/wp0017.pdf (accessed October 8, 2012).

Bonica, Adam. 2012. "Operationalizing the Spatial Model with an Interactive Survey on Public Spending Preferences." *mimeo. Nex York University public files.* https://files.nyu.edu/ajb454/public/files/Bonica_Fisc_Prefs_and_Ideology_3_3.pdf (accessed October 17, 2012).

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Working Papers, Geary Institute, University College Dublin.* 200827.

Bornstein, Gary and Ilan Yaniv. 1998. "Individual and Group Behavior in the Ultimatum Game: Are Groups More 'Rational' Players?" *Experimental Economics* 1:101–108.

Bosch-Domenech, Antoni, Jose Garcia-Montalvo, Rosemarie Nagel, and Albert Satorra. 2002. "One, Two, (Three), Infinity...: Newspaper and Lab Beauty-Contest Experiments." *American Economic Review* 92:1687–1701.

Bottom, W. P., C. L. Eavey, G. J. Miller, and J. N. Victor. 2000. "The institutional effect on majority rule instability: Bicameralism in spatial policy decisions." *American Journal of Political Science* 44:523–540.

Bottom, W. P. and P. W. Paese. 1997. "False consensus, stereotypic cues, and the perception of integrative potential in negotiation." *Journal of Applied Social Psychology* 27:1919–1940.

Bouckaert, J. and G. Dhaene. 2004. "Inter-ethnic trust and reciprocity: results of an experiment with small businessmen." *European Journal of Political Economy* 20:869–886.

Boyd, R. and P. J. Richerson. 1988. "The Evolution of Reciprocity in Sizable Groups." *Journal of Theoretical Biology* 132:337–356.

Bozdogan, H. 1987. "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions." *Psychometrika* 52:345–370.

Bozdogan, H. 2000. "Akaike's Information Criterion and Recent Developments in Information Complexity." *Journal of Mathematical Psychology* 44:62–91.

Bradler, Christiane. 2009. "Social Preferences under Risk - An experimental analysis." *ZEW Discussion Papers, Center for European Economic Research* 09-077.

Bradley, Michael J. 2006. *The Birth of Mathematics: Ancient Times to 1300*. Sonlight Christian.

Bradley, W. J., J. K. Hodge, and D. M. Kilgour. 2005. "Separable discrete preferences." *Mathematical Social Sciences* 49:335–353.

Brady, Henry E. 2008. "Causation and Explanation in Social Science." In *The Oxford Handbook of Political Methodology*, edited by David Collier, Henry E. Brady, and Janet M. Box-Steffensmeier, chapter 10. Oxford: Oxford University Press.

Brambor, Thomas, Roberts Clark William, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analysis." *Political Analysis* 14:63–82.

Brams, S. J., D. M. Kilgour, and W. S. Zwicker. 1997. "Voting on referenda: The separability problem and possible solutions." *Electoral Studies* 16:359–377.

Branas-Garza, Pablo, Miguel A. Duran, and Maria Paz Espinosa. 2009. "The Role of Personal Involvement and Responsibility in Unfair Outcomes : A Classroom Investigation 2009 21: 225." *Rationality and Society* 21:225–248.

Branas-Garza, Pablo, Antonio M. Espin, and Filippos Exadaktylos. 2012. "Students, Volunteers and Subjects: Experiments on Social Preferenes." *Working Papers, Chapman University, Economic Science Institute* 12-11.

Brandenburger, Adam. 1996. "Strategic and Structural Uncertainty in Games." In *Wise Choices: Games, Decisions, and Negotiations*, edited by Richard Zeckhauser, Ralph Keeney, and James Sibenius, pp. 221–232. Harvard Business School Press, Boston.

Brandts, Jordi and Carles Sola. 2006. "Personal Relations and their Effect on Behavior in an Organizational Setting: An Experimental Study." *mimeo. Working paper, Universitat Autonoma de Barcelona.* http://digital.csic.es/bitstream/10261/1698/1/69207.pdf (accessed July 29, 2012).

Braumöller, B. F. 2003. "Causal complexity and the study of politics." *Political Analysis* 11:209–233.

Braumöller, B. F. and G. Görtz. 2000. "The methodology of necessary conditions." *American Journal of Political Science* 44:844–858.

Braun, Dietmar. 2013. "Theorien rationaler Wahl." In *Studienbuch Politikwissenschaft*, edited by M. G. Schmidt, F. Wolf, and S. Wurster. VS Verlag für Sozialwissenschaften, Wiesbaden.

Braunberger, Gerald. 2013. "Wenn Ökonomen streiten: Wie entsteht Wohlstand? Acemoglu/Robinson gegen Jeffrey Sachs." *Frankfurter Allgemeine, Fazit - das Wirtschaftsblog* http://blogs.faz.net/fazit/2013/01/13/wenn-oekonomen-streiten-woher-kommt-der-wohlstand-acemoglu-robinson-versus-jeffrey-sachs-889 (accessed March 17, 2013).

Brennan, Geoffrey and James M. Buchanan. 2000. *The Reason of Rules: Constitutional Political Economy*. Indianapolis, IN: Liberty Fund, Inc.

Brewer, M. 1999. "The Psychology of Prejudice: Ingroup Love or Outgroup Hate?" *Journal of Social Issues* 55:429–444.

Brown, L. D., T. T. Cai, and DasGupta A. 2001. "Interval estimation for a binomial proportion." *Statistical Science* 16:101–133.

Börsch-Supan, Axel and Martin Gasche. 2013. "Sozialsysteme und Demographie - Sägen an den Säulen." *Frankfurter allgemeine Zeitung* 28.01.2013:7.

Bräuninger, Thomas, Andre Bächtiger, and Susumu Shikano. 2012. "Experiment und Simulation." *Jahrbuch für Handlungs- und Entscheidungstheorie* 7.

Brunnhuber, Georg. 2006. "DB-Aufsichtsrat fuer Boersengang." *Eisenbahn-Revue International* 8 – 9:378.

Bräutigam, S. 2005. "Neuroeconomics - From neural systems to economic behaviour." *Brain Research Bulletin* 67:355–360.

Buchanan, James M. 1986. *The Constitution of Economic Policy*. Nobel Prize Lecture.

Buchanan, James M. and Gordon Tullock. 1962. *The calculus of consent: logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.

Bueno de Mesquita, Bruce. 2004. "The methodical study of politics." In *Problems and Methods in the Study of Politics*, edited by Ian Shapiro, Rogers M. Smith, and Tarek E. Masoud, pp. 227–248. Cambridge University Press.

Bueno de Mesquita, Bruce. 2011. "A New Model for Predicting Policy Choices Preliminary Tests." *Conflict Management and Peace Science* 28:65–85.

Buhrmester, M., T. Kwang, and S. D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6:3–5.

Buller, David J. 2005. *Adapting minds : evolutionary psychology and the persistent quest for human nature*. Cambridge, Mass.: MIT Press.

Burgess, Philip M. and James A. Robinson. 1969. "Alliances and the Theory of Collective Action: A Simulation of Coalition Processes." *Midwest Journal of Political Science* 13(2):194–218.

Burkhart, Simone. 2009. "Reforming Federalism in Germany: Incremental Changes instead of the Big Deal." *The Journal of Federalism* 39.

Burnham, Kenneth P. and David R. Anderson. 2004. "Multimodel Inference: Understanding AIC and BIC in Model Selection." In *Sociological Methods Research*, volume 33, pp. 261–304. Sociological Methods and Research.

Butler, D. J. and C. F. Camerer. 2005. "Behavioral game theory: Experiments in strategic interaction." *Journal of Economic Surveys* 19:137–147.

Cabrales, Antonio, Raffaele Miniaci, Marco Piovesan, and Giovanni Ponti. 2010. "Social Preferences and Strategic Uncertainty: An Experiment on Markets and Contracts." *American Economic Review, American Economic Association* 100:2261–2278.

Camerer, C. F. 2003. "Behavioral game theory: Plausible formal models that predict accurately." *Behavioral and Brain Sciences* 26:157ff.

Camerer, Colin F and Ernst Fehr. 2004. "Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists." *Institute for Empirical Research in Economics - University of Zurich, Working Papers* 97:55–95.

Camerer, C. F. and R. M. Hogarth. 1999. "The effects of financial incentives in experiments: A review and capital-labor-production framework." *Journal of Risk and Uncertainty* 19:7–42.

Camerer, C. F., E. J. Johnson, T. Rymon, and S. Sen. 1993. "Cognition and framing in sequential bargaining for gains and losses." In *Frontiers of game theory*, edited by K. Binmore, A. Kirman, and P. Tani, pp. 27–47. Cambridge, MA: MIT Press.

Camerer, C. F., G. Loewenstein, and D. Prelec. 2004. "Neuroeconomics: Why economics needs brains." *Scandinavian Journal of Economics* 106:555–579.

Camerer, C. F. and R. H. Thaler. 1955. "Anomalies: Ultimatums, dictators, and manners." *Journal of Economic Perspectives* 9:209–219.

Cameron, Charles M. and Rebecca Morton. 2002. "Formal Theory Meets Data." In *Political Science: The state of the discipline*, edited by Ira Katznelson and Helen V. Milner, pp. 784–804. London and New York: W.W. Norton and American Political Science Association.

Campbell, Donald T. 1969. "Reforms as experiments." *American Psychologist* 24.

Carmines, Edward G. and Richard A. Zeller. 1979. *Reliability and Validity Assessment*. Quantitative Applications in the Social Science. Sage Publications, London.

Carpenter, Jeffrey and Erika Seki. 2005. "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay." *Discussion paper series, Institute for the Study of Labor* 1697.

Carrubba, C. J. and A. Singh. 2004. "A decision theoretic model of public opinion: Guns, butter, and European common defense." *American Journal of Political Science* 48:218–231.

Carter, John and Michael Irons. 1991. "Are Economists Different, and If So, Why?" *Journal of Economic Perspectives* 5:171–177.

Cason, T. N. and V. L. Mui. 2007. "Communication and coordination in the laboratory collective resistance game." *Experimental Economics* 10:251–267.

Cason, T. N., T. Saijo, T. Yamato, and K. Yokotani. 2004. "Non-excludable public good experiments." *Games and Economic Behavior* 49:81–102.

Castillo, Joan Joseph. 2009. "Convenience Sampling." *Experiment Resources* http://www.experiment-resources.com/convenience-sampling.html (accessed October 1, 2012).

CESS. 2012. "Conducting an experiment at CESS - Rules and procedures associated with conducting an experiment." *Nuffield center for Experimental social science* .

Chajewska, Urszula and Daphne Koller. 1999. *Learning the Structure of Utility Functions*. Stanford: Stanford University.

Chalkiadakis, G. and C. Boutilier. 2004. "Bayesian Reinforcement Learning for Coalition Formation Under Uncertainty." *Autonomous Agents and Multiagent Systems* 04:1090–1097.

Chalkiadakis, G. and C. Boutilier. 2008. "Sequential Decision Making in Repeated Coalition Formation under Uncertainty." *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems* 1:347–354.

Chandrasekhar, Arun G. and Juan Pablo Xandri. 2012. "A note on payments in experiments of infinitely repeated games with discounting." *mimeo. MIT Economics.* http://economics.mit.edu/files/7310 (accessed November 7, 2012).

Charness, G. 2004. "Attribution and reciprocity in an experimental labor market." *Journal of Labor Economics* 22:665–688.

Charness, Gary and Nuno Garoupa. 2000. "Reputation, Honesty, and Efficiency with Insider Information: an Experiment." *Journal of Economics and Management Strategy* 9:425–451.

Charness, Gary and Uri Gneezy. 2008. "What's in a name? Anonymity and social distance in dictator and ultimatum games." *mimeo. Departmental Working Papers, University of California.* http://econ.econ.ucsb.edu/papers/wp11-01.pdf (accessed October 2, 2010).

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. 2012. "Experimental methods: Between-subject and within-subject design." *Journal of Economic Behavior and Organization* 81:1–8.

Charness, G., E. Haruvy, and D. Sonsino. 2007a. "Social distance and reciprocity: An internet experiment." *Journal of Economic Behavior and Organization* 63:88–103.

Charness, G. and M. Rabin. 2002. "Understanding social preferences with simple tests." *Quarterly Journal of Economics* 117:817–869.

Charness, Gary, Luca Rigotti, and Aldo Rustichini. 2007b. "Individual Behavior and Group Membership." *American Economic Association* 97:1340–1352.

Charness, Gary and Aldo Rustichini. 2011. "Gender differences in cooperation with group membership." *Games and Economic Behavior* 72:77–85.

Chen, Y. and S. X. Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99:431–457.

Chen, Y. and P. C. Ordeshook. 1998. "Veto games: Spatial committees under unanimity rule." *Public Choice* 97:617–643.

Chongwoo, Choe and Park In-Uck. 2011. "Information, Authority, and Corporate Hierarchies." *The B.E. Journal of Theoretical Economics* 11.

Chou, E., M. McConnell, R. Nagel, and C. R. Plott. 2009. "The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person "guessing" game." *Experimental Economics* 12:159–179.

Christians, Malte. 2012. "Mehrheit akzeptiert niedrigeres Rentenniveau." *Focus* 41.

Clarke, Kevin A. and David M. Primo. 2007. "Modernizing Political Science: A Model-Based Approach." *Perspectives on Politics* 5:741–753.

Cleave, Blair L., Nikos Nikiforakis, and Robert Slonim. 2010. "Is There Selection Bias in Laboratory Experiments?" *Working Papers University of Sydney, School of Economics* 2123/6957.

Clinton, J. D. and A. Meirowitz. 2004. "Testing explanations of strategic voting in legislatures: A reexamination of the compromise of 1790." *American Journal of Political Science* 48:675–689.

Clopper, C. J. and E. S. Pearson. 1934. "The use of confidence or fiducial limits illustrated in the case of the binomial." *Biometrika* 26:404–413.

Coase, Ronald. 1937. "The Nature of the Firm." *Economica* 4:386–405.

Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3:1–44.

Cohen, Dov and Richard Nisbett. 1997. "Field Experiments Examining the Culture of Honor: The Role of Institutions in Perpetuating Norms about Violence." *Personality and Social Psychology Bulletin* 11:1188–1199.

Coleman, Eric and Elinor Ostrom. 2009. "Experimental Contributions to Collective Action Theory." In *Cambridge Handbook of Experimental Political Science*, edited by Jamie Druckman, Donald Green, Jim Kluklinski, and Arthur Lupia, chapter 24. Cambridge University Press.

Coleman, James. 1990. *Foundations Of Social Theory*. Cambridge: Belkhap Press.

Colman, Andrew M. 2003. "Cooperation, psychological game theory, and limitations of rationality in social interaction." *Behavioral and Brain Sciences* 26:139–198.

Colomer, Josep H. 2001. *Political Institutions: Democracy and Social Choice*. Oxford University Press, New York.

Colomer, Josep M. 1999. "On the Geometry of Unanimity Rule." *Journal of Theoretical Politics* 11:543–553.

Consilium, Council of the European union. 2000. "Council Regulation amending Regulation (EEC) No 404/93 on the common organisation of the market in bananas." *Interinstitutional file - legislative acts and other instruments* 99/0235 (CNS).

Consilium, Rat der europäischen Union. 1999. "Beratungsergebnisse des Vorschlag für eine Verordnung (EG) des Rates zur Stärkung des Dialogs mit den Unternehmen und sonstigen Beteiligten über die Gemeinsame Fischereipolitik." *Interinstitutionelles Dossier* 99/0163 (CNS).

Cooper, D. and K. Kühn. 2009. "Communication, Renegotiation, and the Scope for Collusion." *mimeo. CEPR Discussion Paper, no. 7563.* http://www.cepr.org/pubs/dps/DP7563.asp (accessed July 23, 2012).

Cooper, David J. and John H. Kagel. 2005. "Are Two Heads Better Than One? Team Versus Individual Play in Signaling Games." *American Economic Review* 95:477–509.

Cooper, John O., Timothy E. Heron, and William L. Heward. 2007. *Applied Behavior Analysis*. Pearson, 2nd edition.

Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games:An Experimental Study." *Econometrica* 69:1193–1235.

Coughlin, Peter J. 1990a. "Candidate Uncertainty and Electoral Equilibria." In *Advances in the Spatial Theory of Voting*, edited by James Enelow and Melvin Hinich. NewYork: Cambridge University Press.

Coughlin, Peter J. 1990b. "Majority Rule and Election Models." *Journal of Economic Surveys* 3:157 – 188.

Cox, Gary W. 1984. "Electoral Equilibrium in Double Member Districts." *Public Choice* 44:443–451.

Cox, Gary W. 1997. "Duverger's Law and Strategic Voting." In *Making Votes Count: Strategic Coordination in the World's Electoral Systems*, edited by Gary W. Cox, p. 315. Cambridge University Press.

Cox, James C. 2010. "Some issues of methods, theories, and experimental designs." *Journal of Economic Behavior and Organization* 73:24–28.

Cox, James C. and Cary A. Deck. 2005. "On the Nature of Reciprocal Motives." *Economic Inquiry, Western Economic Association International* 43:623–635.

Cox, James C. and Glenn W. Harrison. 2008. *Risk Aversion in Experiments*, volume 12 of *Research in Experimental Economics Series*. Emerald Group Publishing.

Cox, James C., Klarita Sadiraj, and Vjollca Vjollca. 2008. "Implications of Trust, Fear, and Reciprocity for Modeling Economic Behavior." *Experimental Economics* 11:1–24.

Cox, James C. and Vjollca Sadiraj. 2012. "Direct Tests Of Individual Preferences For Efficiency And Equity." *Economic Inquiry, Western Economic Association International* 50:920–931.

Craig, Stephen, David Hill, and Scott Richards. 2006. "Citizen Attitudes about Divided Government in the United States: Does the Public Really Care?" *mimeo. Paper presented at the annual meeting of the Southern Political Science Association.* http://www.allacademic.com/meta/p68426_index.html (accessed July 16, 2012).

Crawford, Vincent P. 1985. "Dynamic Games and Dynamic Contract Theory." *The Journal of Conflict Resolution* 29:195–224.

Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. 2010. "Strategic thinking." *mimeo. University of California.* http://dss.ucsd.edu/~vcrawfor/CGCI27Dec10.pdf (accessed April 8, 2011).

Crawford, Vincent P. and Hans Haller. 1990. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." *Econometrica* 58:571–595.

Croson, R. 1999. "Handbook of experimental economics." *Journal of Economic Behavior and Organization* 40:115–118.

Croson, Rachel. 2005. "The Method of Experimental Economics." *International Negotiation* 10:131–148.

Croson, R. and U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47:448–474.

Cross, J. P. 2012. "Interventions and negotiation in the Council of Ministers of the European Union." *European Union Politics* 13:47–69.

Croux, C., G. Dhaene, and D. Hoorelbeke. 2003. "Robust Standard Errors for Robust Estimators. Research." *Discussions Paper Series, Department of Applied Economics, Leuven, Belgium.* 03.16.

Dal Bo, Pedro, Andrew Foster, and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review* 100:2205–2229.

Dalio, Ray. 2011. "A Template for Understanding - How the Economic Machine Works and How it is Reflected Now." *mimeo. Bridgewater Associates.* http://www.zerohedge.com/sites/default/files/images/user5/imageroot/2011/12/Ray%20Dalio%27s%20Template%20for%20Understanding%20%28Oct-2011%29.pdf (accessed December 2, 2012).

Dalton, Russell J., Wilhelm Bürklin, and Andrew Drummond. 2001. "Public opinion and direct democracy." *Journal of Democracy* 12.

Dana, J., R. A. Weber, and J. X. Kuang. 2007. "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness." *Economic Theory* 33:67–80.

Davis, Douglas D. and Charles A. Holt. 1993. *Experimental economics*. Princeton, N.J.: Princeton University Press.

Davis, O. and M. Hinich. 1966. "A mathematical model of policy formation in a democratic society." In *Mathematical applications in political science*, edited by J. Bernd, volume II, pp. 175–208. Southern Methodist University Press, Dallas.

Davis, O. and M. Hinich. 1967. "Some results related to a mathematical model of policy formation in a democratic society." In *Mathematical applications in political science*, edited by J. Bernd, volume III, pp. 14–38. University Press of Virginia, Charlottesville.

Davis, O. and M. Hinich. 1968. "On the power and importance of the mean preference in a mathematical model of democratic choice." *Public Choice* 5:59–72.

Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *The American Political Science Review* 64:426 – 448.

Dawes, Robyn M., John M. Orbell, Randy T. Simmons, and Alphons J. C. Van De Kragt. 1986. "Organizing Groups for Collective Action." *The American Political Science Review* 80(4):1171–1185.

DBT, Deutscher Bundestag. 2011. "Gesetzentwurf Bundesregierung: Feststellung des Bundeshaushaltsplans fuer das Haushaltsjahr 2011 (Haushaltsgesetz 2011)." *Bundestagsdrucksachen* 17/2500.

DBT, Deutscher Bundestag. 2012. "Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Dorothea Steiner, Nicole Maisch, Baerbel Hoehn, weiterer Abgeordneter und der Fraktion BUENDNIS 90/DIE GRUENEN: Weiterentwicklung der REACH-Verordnung - Position der Bundesregierung in der Europaeischen Debatte." *Bundestags Drucksache* 17/9524.

de Groot-van Leeuwen, Leny E. and Wannes Rombouts. 2010. *Separation of Powers in Theory and Practice - An International Perspective*. Wolf Legal Publishers.

DellaVigna, Stefano. 2007. "Psychology and Economics: Evidence from the Field." *NBER Working Paper Series* 13420.

Demertzis, Maria. 2009. "The 'Wisdom of the Crowds' and Public Policy." *DNB Working Papers, Netherlands Central Bank* 203.

Denzau, Arthur T. and Robert J. Mackay. 1981. "Structure-Induced Equilibria and Perfect-Foresight Expectations." *American Journal of Political Science* 25:762–779.

Dewan, Torun and Rafael Hortala-Vallve. 2009. "The Three A's of Government Formation: Appointment, Allocation, and Assignment." *PSPE working papers* p. 34.

Diederich, Johannes and Timo Goeschl. 2011. "Giving in a Large Economy: Price vs. Non-Price Effects in a Field Experiment." *Discussion Paper Series, Department of Economics, University of Heidelberg* 514.

Diermeier, Daniel and Sean Gailmard. 2006. "Self-Interest, Inequality, and Entitlement in Majoritarian Decision-Making." *Quarterly Journal of Political Science* 1:327–350.

Diermeier, Daniel and Rebecca Morton. 2005. "Experiments in Majoritarian Bargaining." In *Social Choice and Strategic Decisions*, edited by David Austen-Smith and John Duggan, pp. 201–236. Berlin and New York: Springer: Springer.

Dittrich, Dennis A.V. and Martin G. Kocher. 2011. "Monitoring and Pay: An Experiment on Employee Performance under Endogenous Supervision." *mimeo. Papers on Strategic Interaction, Max Planck Institute of Economics, Strategic Interaction Group.* 2006-23. http://epub.ub.uni-muenchen.de/12222/1/Dittrich_Kocher_2011_mimeo_Monitoring_and_Pay.pdf (accessed August 11, 2012).

Dittrich, Dennis A. and Anthony Ziegelmayer. 2010. "Laboratory bilateral gift exchange: Heterogeneous social preferences." *mimeo. Jacobs University Bremen.* http://economicscience.net/content/publications (accessed August 11, 2012).

Dixit, Avinash K., Susan Skeath, and David Reiley. 2009. *Games of strategy*. New York: W. W. Norton & Co.

Djorgovski, George. 2005. "Virtual Astronomy, Information Technology, and the New Scientific Methodology." *Seventh International Workshop on Computer Architectures for Machine Perception (CAMP 2005), 4-6 July 2005, Palermo, Italy* pp. 125–132.

Dockner, E. J. and N. V. Long. 1993. "International Pollution-Control - Cooperative Versus Noncooperative Strategies." *Journal of Environmental Economics and Management* 25:13–29.

Doina, Bratu Renate and Petria Nicolae. 2011. "The Structure of Financial Regulation." *Ovidius University Annals, Economic Sciences Series, Ovidius University of Constantza, Faculty of Economic Sciences* 0:128–132.

Domingos, Pedro. 1999. "The role of Occam's Razor in knowledge discovery." *Data Mining and Knowledge Discovery* 3:409–425.

Dopfner, Kurt. 2002. "Die Rückkehr des verlorenen Menschen - Wege zum Homo sapiens oeconomicus." In *Psychologische Grundlagen der Ökonomie*, edited by Fehr, pp. 99–106. Neue Zürcher Zeitung.

Dougherty, Keith, Brian Pitts, Justin Moeller, and Robi Ragan. 2009. "An Experimental Study of the Efficiency of Unanimity Rule and Majority Rule." *mimeo. Indiana University Bloomington.* http://www.indiana.edu/~wow4/papers/dougherty_pitts_wow4.pdf (accessed January 27, 2012).

Dougherty, Keith L. and Julian Edward. 2004. "The Pareto Efficiency and Expected Costs of k-Majority Rules." *Politics Philosophy and Economics* 3:161–89.

Dougherty, Keith L. and Julian Edward. 2012. "Voting for Pareto optimality: a multidimensional analysis." *Public Choice* 151:655–678.

Dow, Alexander and Sheila C. Dow. 2011. "Animal Spirits Revisited." *Capitalism and Society* 6:1. Berkeley Electronic Press.

Dowding, Keith and Desmond King. 1995. *Introduction to Preferences, Institutions, and Rational Choice*. Oxford: Clarendon Press.

Downs, Anthony. 1957. *An economic theory of democracy*. New York,: Harper.

Downs, Anthony and Rudolf Wildenmann. 1968. *Oekonomische Theorie der Demokratie*. Tuebingen,: Mohr.

DPA. 2011a. "Bahn: Boersengangs-Diskussion momentan ueberfluessig." *Verkehrs Rundschau* 02.03.

DPA. 2011b. "Finanzkrise stoppt Boersengang der Bahn." *Die Zeit* 14.11.

Draper, Janet. 2004. "The relationship between research question and research design." In *Research into Practice: Essential Skills for Reading and Applying Reasearch in Nursing and Health Care*, edited by Patrick A. Crookes and Sue Davies, chapter 6, pp. 69–84. Edinburgh: Bailliere Tindall.

Druckman, James and Arthur Lupia. 2006. "Mind, Will, and Choice." In *The Oxford Handbook of Contextual Political Analysis*, edited by Joseph L. Tilly, Charles and Robert E Goodin, chapter 5. Oxford University Press.

Druckman, J. N. 2004. "Political preference formation: Competition, deliberation, and the (Ir)relevance of framing effects." *American Political Science Review* 98:671–686.

Druckman, James N. 2011. *Cambridge Handbook of Experimental Political Science*. Cambridge ; New York: Cambridge University Press.

Druckman, James N. and Cindy D. Kam. 2009. "Students as Experimental Participants: A Defense of the narrow Data Base." In *Cambridge Handbook of Experimental Political Science*, edited by J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia. Boston: Cambridge University Press.

Druckman, J. N. and M. Parkin. 2005. "The impact of media bias: How editorial slant affects voters." *Journal of Politics* 67:1030–1049.

Dufwenberg, Martin, Paul Heidhues, Georg Kirchsteiger, Frank Riedel, and Joel Sobel. 2008. "Other-Regarding Preferences in General Equilibrium." *CEPR Discussion Papers* 6815.

Dutcher, E. Glenn. 2012. "The effects of telecommuting: An experimental examination. The role of dull and creative tasks." *Journal of Economic Behavior and Organization* 84:355–363.

Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modem State.* New York: Wiley.

Dyer, J. S., P. C. Fishburn, R. E. Steuer, J. Wallenius, and S. Zionts. 1992. "Multiple Criteria Decision-Making, Multiattribute Utility-Theory - the Next 10 Years." *Management Science* 38:645–654.

Easton, David. 1957. "An Approach to the analysis of Political systems." *World Politics* 9:383–400.

Eavey, Cheryl L. and Gary J. Miller. 1984. "Fairness in Majority Rule Games with a Core." *American Journal of Political Science* 28:570–586.

Eavey, Cheryl L. and Gary J. Miller. 1995. "Subcommittee Agenda Control." *Journal of Theoretical Politics* 7:125–156.

EC, European Commission. 2000. "Commission proposes solution to end banana dispute." *mimeo. Communication of the European Commission to the European Council, IP/00/1110.* http://europa.eu/rapid/press-release_IP-00-1110_en.htm?locale=en (accessed February 9, 2012).

Eckel, Catherine and Rick Wilson. 2006. "Internet Cautions: Experimental Games with Internet Partners." *Experimental Economics* 9:53–66.

Eckel, Catherine C. and Philip J. Grossman. 1996. "Altruism in Anonymous Dictator Games." *Games and Economic Behavior* 16:181–191.

Eckel, Catherine C. and Philip J. Grossman. 1998. "Are Women Less Selfish Than Men? Evidence from Dictator Experiments." *Economic Journal* 108:726–735.

Eckel, Catherine C. and Ragan Petrie. 2011. "Face Value." *American Economic Review* 101:1497–1513.

Edwards, W. and J.R Newman. 1982. *Multiattribute evaluation*. Beverly Hills, CA: Sage.

Eeckhoudt, Louis, Christian Gollier, and Harris Schlesinger. 2005. *Economic and Financial Decisions under Risk*. Princeton University Press.

Efron, Bradley and Robert Tibshirani. 1993. *An introduction to the bootstrap*. New York: Chapman & Hall.

Ehmke, Mariah D., Jayson L. Lusk, and Wallace E. Tyner. 2005. "A Multidimensional Homo Economicus: Cultural Dimensions of Economic Preferences in Four Countries." *Presented at the Annual meeting of the Agricultural and Applied Economics Association, July 24-27* .

Ehrhart, Karl-Martin, Roy Gardner, Juergen von Hagen, and Claudia Keser. 2007. "Budget processes: Theory and experimental evidence." *Games and Economic Behavior* 59:279–295.

Eichberger, Jürgen and David Kelsey. 2008. "Ambiguity." In *The Oxford Handbook of Rational and Social Choice*, edited by P. Anand, P. Pattanaik, and C. Puppe, chapter V, pp. 113–139. Oxford: Oxford University Press.

Eisenkopf, Gerald and Sabrina Teyssier. 2010. "Envy and Loss Aversion in Tournaments." *Research Paper Series, Thurgau Institute of Economics and Department of Economics (University of Konstanz)* 52.

El-Gamal, Mahmoud A. and David Grether. 1995. "Are People Bayesian? Uncovering Behavioral Strategies." *Journal of the American Statistical Association: Applications and Case Studies* 90:1137–1145.

Ellsberg, D. 1961. "Risk, ambiguity and the Savage axioms." *Quarterly Journal of Economics* 75:643–669.

Elster, Jon. 1998. *Deliberative Democracy*. Cambridge: Cambridge University Press.

Enelow, James M. and Melvin J. Hinich. 1984. *The spatial theory of voting : an introduction*. Cambridge Cambridgeshire ; New York: Cambridge University Press.

Engelmann, D. and M. Strobel. 2000. "An Experimental Comparison of the Fairness Models by Bolton and Ockenfels and by Fehr and Schmidt." *Schriftenreihe Sonderforschungsbereich 373, Humboldt Universitaet Berlin.* 28.

Engelmann, D. and M. Strobel. 2004. "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments." *American Economic Review* 94:857–869.

EP, European Parliament. 2005. "REACH - MEPs debate chemicals legislation and new chemicals agency environment." http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+IM-PRESS+20051116IPR02381+0+DOC+XML+V0//EN (accessed October 11, 2011).

Epanechnikov, V. A. 1969. "Nonparametric estimation of a multidimensional probability density." *Theory of Probability and Its Applications* 14:153–158.

Epstein, David and Sharyn O'Halloran. 1996. "Divided Government and the Design of Administrative Procedures: A Formal Model and Empirical Test." *The Journal of Politics, Cambridge University Press* 58:373–397.

Epstein, David and Sharyn O'Halloran. 1999. *Delegating powers: a transaction cost politics approach to policy making under separate powers*. Cambridge, U.K. ; New York: Cambridge University Press.

Eriksson, K. and B. Simpson. 2010. "Emotional reactions to losing explain gender differences in entering a risky lottery." *Judgment and Decision Making* 5:159–163.

Erlei, Mathias. 2004. "Heterogeneous Social Preferences." *Working Papers in Economics, Technische Universitaet Clausthal* 0001.

Erlenmaier, Ulrich and Hans Gersbach. 2001. "Flexible Majority Rules." *CESifo Working Paper Series* 464.

Ert, Eyal and Ido Erev. 2008. "The rejection of attractive gambles, loss aversion, and the lemon avoidance heuristic." *Journal of Economic Psychology* 29:715–723.

Ert, Eyal, Ido Erev, and Alvin E. Roth. 2011. "A Choice Prediction Competition for Social Preferences in Simple Extensive Form Games: An Introduction." *Games, MDPI.* 2:257–276.

Evans, J. 2003. "In two minds: dual-process accounts of reasoning." *TRENDS in Cognitive Sciences* 7.

Experimental-Psychology. 2012. "The roar of the crowd - Crowdsourcing is transforming the science of psychology." *The Economist* print edition May 26th:69–70.

Faas, Thorsten and Sascha Huber. 2010. "Experimente in der Politikwissenschaft: Vom Mauerbluemchen zum Mainstream." *Politische Vierteljahresschrift (PVS)* 51:721–749.

Falk, Armin. 2009. "Lab Experiments Are a Major Source of Knowledge in the Social." *Science* 326:535–538.

Falk, Armin, Stephan Meier, and Christian Zehnder. 2011. "Do Lab Experiments Misrepresent Social Preferences? The case of self-selected student samples." *Journal of the European Economic Association* 11.

FDP. 2007. "Echter Wettbewerb fuer mehr Verkehr auf der Schiene." *Beschluss des FDP-Praesidiums* Pressemitteilung vom 24.09.2007, 13:06.

Feddersen, T., S. Gailmard, and A. Sandroni. 2009. "Moral Bias in Large Elections: Theory and Experimental Evidence." *American Political Science Review* 103:175 192.

Feddersen, Timothy J. 1992. "A Voting Model Implying Duverger's Law and Positive Turnout." *American Journal of Political Science* 36:938–962.

Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature* 454:1079–1084.

Fehr, E. and U. Fischbacher. 2002. "Why social preferences matter - The impact of non-selfish motives on competition, cooperation and incentives." *Economic Journal* 112:C1–C33.

Fehr, Ernst and Urs Fischbacher. 2003. "The nature of human altruism." *Nature* 425:785–791.

Fehr, E., U. Fischbacher, and S. Gachter. 2002a. "Strong reciprocity, human cooperation, and the enforcement of social norms." *Human Nature-an Interdisciplinary Biosocial Perspective* 13:1–25.

Fehr, Ernst, Alexander Klein, and Klaus M. Schmidt. 2001. "Fairness, Incentives and Contractual Incompleteness." *CEPR Discussion Papers* 2790.

Fehr, E., S. Kremhelmer, and K.M. Schmidt. 2002b. "Fairness and the Optimal Allocation of Ownership Rights." *Discussion Paper Series, Governance and the Efficiency of Economic Systems* 11. Free University of Berlin, Humboldt University of Berlin, University of Bonn, University of Mannheim, University of Munich.

Fehr, E. and K. M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics* 114:817–868.

Fehr, Ernst and Klaus M. Schmidt. 2000. "Theories of Fairness and Reciprocity - Evidence and Economic Applications." *Working Papers, Institute for Empirical Research in Economics, University of Zurich.* iewwp075.

Fehr, Ernst and Gerhard Schwarz. 2002. *Psychologische Grundlagen der Ökonomie*. Neue Zürcher Zeitung.

Feinerer, Ingo, Kurt Hornik, and David Meyer. 2008. "Text Mining Infrastructure in R." *Journal of Statistical Software* 25:1–54.

Felsenthal, Dan, Amnon Rapoport, and Zeev Maoz. 1988. "Tacit cooperation in three alternative noncoopera tive voting games: a new model of sophisticated behavior under the plurality procedure." *Electoral Studies* 7:143–61.

Ferejohn, John A. and Morris P. Fiorina. 1975. "Purposive Models of Legislative Behavior." *American Economic Review* 65:407–414.

Fernandez, Raquel and Dani Rodrik. 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81:1146–1155.

Ferris, Gerald R. and Wayne A. Hochwarter. 2011. "Organizational politics." In *Maintaining, expanding, and contracting the organization*, edited by Sheldon Zedeck, volume 3 of *APA Handbooks in Psychology; APA handbook of industrial and organizational psychology*, pp. 435–459. Washington, DC, US: American Psychological Association.

Fey, Mark. 1997. "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting." *The American Political Science Review* 91:135–147.

Fey, M. 2004. "May's Theorem with an infinite population." *Social Choice and Welfare* 23:275–293.

Fiedler, Marina and Ernan Haruvy. 2009. "The Lab versus the Virtual Lab and Virtual Field - An Experimental Investigation of Trust Games with Communication." *Journal of Economic Behavior and Organization* 72:716–724.

Finke, D. 2009. "Estimating the Effect of Nonseparable Preferences in Eu Treaty Negotiations." *Journal of Theoretical Politics* 21:543–569.

Finke, Daniel and Andreas Fleig. 2013. "The Merits of Adding Complexity: Non-Separable Preferences in Spatial Models of EU Politics." *Journal of Theoretical Politics* 25:547 – 576.

Fiorina, Morris P. 1992a. *Divided Government*. Allyn and Bacon, Boston.

Fiorina, Morris P. 1992b. "An Era of Divided Government." *Political Science Quarterly* 107:387–410.

Fiorina, Morris P. and Charles R. Plott. 1978. "Committee Decisions under Majority Rule: An Experimental Study." *The American Political Science Review* 72:575–598.

Fiorino, Nadia and Roberto Ricciuti. 2007. "Determinants of Direct Democracy." *International Centre for Economic Research (ICER) Working Papers* 23 - 2007.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10:171–178.

Fischbacher, U., A. Falk, and E. Fehr. 2008. "Testing theories of fairness - Intentions matter." *Games and Economic Behavior* 62:287–303.

Fischbacher, U., S. Gachter, and E. Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters* 71:397–404.

Fishburn, P. 1977. "Condorcet Social Choice Functions." *SIAM Journal of Applied Mathematics* 33:469–489.

Fishburn, Peter C. 1973. *The theory of social choice*. Princeton, N.J.: Princeton University Press.

Fleig, Andreas and Daniel Finke. 2013. "Delegation, Uncertainty and Social Preferences in Majority Decisions?" mimeo.

Forsa. 2011. "Würden Sie für die Griechen zahlen?" *Statista* http://de.statista.com/statistik/daten/studie/201617/umfrage/zahlungsbereitschaft-der-deutschen-fuer-griechenland (accessed August 3, 2012).

Forsythe, R., T. Rietz, R. Myerson, and Weber R. 1996. "An experimental study of voting rules and polls in three-candidate elections." *International Journal of Game Theory* 25:355–383.

Frank, Bjoern. 1998. "Good news for experimenters: subjects do not care about your welfare." *Economics Letters* 61:171–174.

Frank, Robert, Thomas Gilovich, and Dennis Regan. 1993. "Does studying economics inhibit cooperation?" *Journal of Economic Perspectives* 7:159–171.

Frank, Robert, Thomas Gilovich, and Dennis Regan. 1996. "Do economists make bad citizens?" *Journal of Economic Perspectives* 10:187–192.

Franz, Stephan. 2004. "Grundlagen des oekonomischen Ansatzes: das Erklaerungskonzept des Homo Oeconomicus." *International Economics, Working Paper Uni Potsdam* 02.

Frechette, Guillaume R. 2011. "Laboratory Experiments: Professionals Versus Students." *mimeo. New York University, working paper.* http://ssrn.com/abstract=1939219 (accessed March 5, 2012).

Frechette, Guillaume R. 2012. "Session-effects in the laboratory." *Experimental Economics* 15:485–498.

Frei, Brent. 2009. "Paid Crowdsourcing: Current State & Progress toward Main-stream Business Use." *mimeo. Smartsheet.com.* http://www.smartsheet.com/files/haymaker/Paid%20Crowdsourcing%20Sept%20 2009%20-%20Release%20Version%20-%20Smartsheet.pdf (accessed July 19, 2012).

Frevert, Ute. 2001. *Die kasernierte Nation. Militaerdienst und Zivilgesellschaft in Deutschland.* Verlag C. H. Beck, Muenchen.

Frey, Bruno S. and Stephan Meier. 2005. "Selfish and Indoctrinated Economists?" *European Journal of Law and Economics* 19:165–171.

Friedman, Thomas. 2007. "The World is Flat 3.0." *Speech at MIT, December 2nd.* http://www.academicearth.org/lectures/the-world-is-flat (accessed July 17, 2012).

Frohlich, Norman and Joe A. Oppenheimer. 1996. "Experiencing Impartiality to Invoke Fairness in the n-PD: Some Experimental Results." *Public Choice* 86:117–135.

Frohlich, Norman and Joe A. Oppenheimer. 2007. "Justice preferences and the Arrow problem." *Journal of Theoretical Politics* 19:363–390.

Gailmard, S. 2009. "Discretion Rather than Rules: Choice of Instruments to Control Bureaucratic Policy Making." *Political Analysis* 17:25–44.

Ganghof, Steffen and Philip Manow. 2005. *Mechanismen der Politik: Strategische Interaktion im deutschen Regierungssystem*, volume 54 of *Schriften des Max-Planck-Instituts fuer Gesellschaftsforschung*. Frankfurt a.M.: Campus.

Gautier, Axel and Dimitri Paolini. 2007. "Delegation and information revelation." *Journal of Institutional and Theoretical Economics* 163:574–597.

Geiger, David, Michael Rosemann, Erwin Fielt, and Martin Schader. 2012. "Crowdsourcing Information Systems - Definition Typology, and Design." *mimeo. Proceedings of the International Conference on Information Systems.* https://ub-madoc.bib.uni-mannheim.de/32631/1/Geiger_et_al._-_2012_-_Crowdsourcing_Information_Systems_-_Definition,_Typology,_and_Design.pdf (accessed December 1, 2012).

Geisendorf, Sylvie. 2009. "The irrational foundations of neoclassical economics. How a simplified Newtonism shaped the perception of the economy." In *Der neoliberale Markt-Diskurs: Ursprünge, Geschichte, Wirkungen*, edited by Walter Otto Ötsch and Claus Thomasberger, pp. 151–166. Metropolis.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal* 1:2:35–52.

Gerring, John. 2005. "Causation - a unified framework for the social sciences." *Journal of Theoretical Politics* 17:163–198.

Gibbons, J.D. 1976. *Nonparametric Methods for Quantitative Analysis*. Holt, Rhinehart, and Winston.

Gigerenzer, G. 2008. *Rationality for mortals: How people cope with uncertainty*. Oxford.

Gigerenzer, G. and W. Gaissmaier. 2011. "Heuristic decision making." *Annual Review of Psychology* 62:451–482.

Gigerenzer, Gerd and Daniel G. Goldstein. 1999. "Betting on One Good Reason: The Take The Best Heuristic." In *Simple heuristics that make us smart*, edited by Research Group ABC, chapter 4. Oxford University Press, New York.

Gigerenzer, G. and R. Selten. 2001. *Bounded rationality: The adaptive toolbox*. Cambridge, Mass.

Gilboa, Itzhak. 2010. *Rational Choice*, volume 1. The MIT Press, 1 edition.

Gill, Jeff. 2006. *Essential mathematics for political and social research*. Cambridge; New York: Cambridge University Press.

Gillespie, Richard. 1991. *Manufacturing knowledge : a history of the Hawthorne experiments*. Cambridge England ; New York: Cambridge University Press.

Gillies, D. B. 1959. "Solutions to general non-zero-sum games." In *Contributions to the Theory of Games*, edited by A. W. Tucker and R. D. Luce, volume IV, pp. 47–85. Princeton: Princeton University Press.

Gilligan, Thomas W. and Keith. Krehbiel. 1987. "Collective Decision-Making and Standing Committees: An Informational Rational for Restrictive Amendment Procedures." *Journal of Law, Economics, and Organization* 3:287–334.

Gilligan, T. W. and K. Krehbiel. 1990. "Organization of Informative Committees by a Rational Legislature." *American Journal of Political Science* 34:531–564.

Gini, C. 1912. *Variabilità e mutabilità*. Bologna, Italy.

Gladwell, Malcolm. 2007. *Blink: The Power of Thinking Without Thinking*. Back Bay Books.

Glasgow, Garrett, Matt Golder, and Sona N. Golder. 2012. "New Empirical Strategies for the Study of Parliamentary Government Formation." *Political Analysis* 20:248–270.

Gneezy, U. and J. Potters. 1997. "An experiment on risk taking and evaluation periods." *The Quarterly Journal of Economics* 112:631–645.

Goeree, J. K. and C. A. Holt. 2001. "Ten little treasures of game theory and ten intuitive contradictions." *American Economic Review* 91:1402–1422.

Goeree, Jacob K., Charles A. Holt, and Susan K. Laury. 2002. "Private costs and public benefits: unraveling the effects of altruism and noisy behavior." *Journal of Public Economics* 83:255–276.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey. 2005. "Regular Quantal Response Equilibrium." *Experimental Economics* 8:347–367.

Goerg, Sebastian J., Werner Güth, Gari Walkowitz, and Torsten Weiland. 2007. "Interregional diversity of fairness concerns - An online ultimatum experiment." *Jena Economic Research Papers, Max-Planck-Institute of Economics* 016.

Goette, Lorenz, David Huffman, and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups." *American Economic Review* 96:212–216.

Goode, W. 1971. "Force and Violence in the Family." *Journal of Marriage and the Family* 33:624–636.

Gorman, W. M. 1968. "The Structure of Utility Functions." *The Review of Economic Studies* 35:367–390.

Grant, robert M. 1996. "Toward a knowledge-based theory of the firm." *Strategic Management Journal* 17:109–122.

Gray, Jeremy. 2011. "The vagueness of Wigner's analysis." *Interdisciplinary Science Reviews* 36:214–228.

Green, Donald P. and Ian Shapiro. 1994. *Pathologies of Rational Choice Theory. A Critique of Applications in Political Science*. New Haven-London: Yale University Press.

Greene, W.H. 2003. *Econometric analysis*. New Jersey: Prentice Hall.

Greiner, Ben. 2004. "An Online Recruitment System for Economic Experiments." In *Forschung und wissenschaftliches Rechnen*, edited by Kurt Kremer and Volker Macho, volume 63 of *GWDG Bericht*, pp. 79–93. Goettingen: Gesellschaft fuer wissenschaftliche Datenverarbeitung.

Grendstad, Gunnar and Per Selle. 1995. "Cultural Theory and the New Institutionalism." *Journal of Theoretical Politics* 7:5–27.

Grether, D. M. and C. R. Plott. 1979. "Economic Theory of Choice and the Preference Reversal Phenomenon." *American Economic Review* 69:623–638.

Gräfe, A. and J. S. Armstrong. 2010. "Predicting Elections from the Most Important Issue: A Test of the Take-the-best Heuristic." *Journal of Behavioral Decision Making* 25:41–48.

Grosslight, Lorraine, Christopher Unger, Eileen Jay, and Carol L. Smith. 1991. "Understanding models and their use in science: Conceptions of middle and high school students and experts." *Journal of Research in Science Teaching, Special Issue: Students' Models and Epistemologies* 28:799–822.

Güth, Werner, M. Vittoria Levati, Matthias Sutter, and Eline van der Heijden. 2004. "Leadership and cooperation in public goods experiments." *Papers on Strategic Interaction, Max Planck Institute of Economics* 29.

Güth, Werner, Matteo Ploner, and Vittoria Levati. 2005. "The Effect of Group Identity in an Investment Game." *Papers on Strategic Interaction. Max Planck Institute of Economics* 06.

Güth, W., R. Schmittberger, and B. Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journalof Economic Behavior and Organization* 3:367–388.

Guala, F. 2005. *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press.

Guarnaschelli, Serena, Richard D. McKelvey, and Thomas R. Palfrey. 2000. "An Experimental Study of Jury Decision Rules." *The American Political Science Review* 94:407–423.

Gujarati, Damodar and Dawn Porter. 2008. *Basic Econometrics*. Mcgraw-Hill Higher Education, 4th edition.

Gul, Faruk. 1997. "A Nobel Prize for Game Theorists: The Contributions of Harsanyi, Nash and Selten." *Journal of Economic Perspectives, American Economic Association* 11:159–174.

Gutmann, Amy and Dennis Thompson. 2004. *Why Deliberative Democracy?* Oxford: Princeton University Press.

Guttman, L. 1945. “A basis for analyzing test-retest reliability.” *Psychometrika* 10:255–282.

Haege, Frank M. 2007. “Committee Decision-making in the Council of the European Union.” *European Union Politics* 8:299–328.

Hagemann, S. 2007. “Applying ideal point estimation methods to the council of ministers.” *European Union Politics* 8:279–296.

Hagemann, Sara and Julia De Clerck-Sachsse. 2007. “Decision-Making in the Enlarged Council of Ministers: Evaluating the Facts.” *CEPS - Policy brief* 119.

Haile, Philip A., Ali Hortacsu, and Grigory Kosenok. 2008. “On the Empirical Content of Quantal Response Equilibrium.” *The American Economic Review* 98:180–200.

Halevy, A., P. Norvig, and F. Pereira. 2009. “The Unreasonable Effectiveness of Data.” *IEEE Intelligent Systems* 24:8–12.

Halfpenny, Peter and Michael Taylor. 1973. “An Experimental Study of Individual and Collective Decision-Making.” *British Journal of Political Science* 3:425–444.

Hall, Peter A. and Rosemary C. R. Taylor. 1996. “Political Science and the Three New Institutionalisms.” *Political Studies* XLIV:936–957.

Ham, J. C., J. H. Kagel, and S. F. Lehrer. 2005. “Randomization, endogeneity and laboratory experiments: the role of cash balances in private value auctions.” *Journal of Econometrics* 125:175–205.

Hamenstädt, Ulrich. 2012. “Das "Lab in the field" Experiment. Kontrolle und die Integration finanzieller Anreize in feldexperimentellen Forschungsdesigns anhand eines Beispiels.” In *Jahrbuch der Handlungs- und Entscheidungstheorie*, edited by Thomas Bräuninger, Andre Bächtiger, and Susumu Shikano, number Band 7, chapter 5, pp. 105–124. VS Verlag für Sozialwissenschaften, Wiesbaden.

Hamming, R. W. 1980. “The Unreasonable Effectiveness of Mathematics.” *The American Mathematical Monthly* 87:81–90.

Hammon, L. and H. Hippner. 2012. “Crowdsourcing.” *Business and Information Systems Engineering* 4.

Hammond, Thomas H. and Gary J. Miller. 1987. “The Core of the Constitution.” *The American Political Science Review* 81:1155–1174.

Hansen, John Mark. 1998. “Individuals, Institutions, and Pub-lic Preferences over Public Finance.” *American Political Sci-ence Review* 92:513–532.

Harrison, Glenn and John List. 2004. “Field Experiments.” *Journal of Economic Literature* 42:1009–1055.

Harrison, Glenn W. 2008. “Maximum Likelihood Estimation of Utility Functions Using Stata.” *Working Paper, Department of Economics, University of Central Florida* 06-12.

Harrison, Glenn W. and E. Elisabet Rutstroem. 2008. "Risk Aversion in the Laboratory." In *Risk Aversion in Experiments*, edited by Glenn W. Harrison and James C. Cox, volume 12 of *Research in Experimental Economics*, chapter 3, pp. 41–196. Emerald Group Publishing.

Hastie, Reid and Robyn M. Dawes. 2010. *Rational choice in an uncertain world : the psychology of judgment and decision making*, volume 2. Los Angeles: SAGE.

Haughton, Jonathan and Shahidur R. Khandker. 2009. *Handbook on Poverty and Inequality, by*. Washington D.C.: The world Bank.

Hayes-Renshaw, Fiona, Wim van Aken, and Helen Wallace. 2006. "When and Why the EU Council of Ministers Votes Explicitly." *Journal of Common Market Studies* 44:1468–5965.

Hayes-Renshaw, Fiona and Hellen Wallace. 2006. *The Council of Ministers*. London: Macmillan, 2nd edition.

Heilbronner, Robert. 2000. *The Worldly Philosophers: The Lives, Times and Ideas of the Great Economic Thinkers*. Penguin Books, 7 edition.

Heinemann, Frank. 2005. "Measuring Risk Aversion and the Wealth Effect." In *Risk Aversion in Experiments - Research in Experimental Economics*, edited by James C. Cox and Glenn W. Harrison, number 12, pp. 293–313. Bingley, UK: Emerald.

Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels. 2009. "Measuring Strategic Uncertainty in Coordination Games." *Review of Economic Studies* 76:181–221.

Heinrich, Gert. 2006. *Grundlagen der Mathematik, der Statistik und des Operations Research für Wirtschaftswissenschaftler [Taschenbuch]*. Oldenbourg Wissenschaftsverlag, 3. auflage edition.

Hellman, Ziv. 2008. "Bargaining Set Solution Concepts in Dynamic Cooperative Games." *MPRA Paper, University Library of Munich* 8798.

Helson, H. 1964. *Adaptation level theory: An experimental and systematic approach to behavior*. Harper and Row.

Hennig-Schmidt, Heike. 1997. "Break-offs in Bargaining, Evidence from a Video Experiment." *Discussion Paper Serie, University of Bonn* B - 402.

Henrich, Joseph, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review* 91:73–78.

Henrich, J., S. J. Heine, and A. Norenzayan. 2010. "The weirdest people in the world?" *Behavioral and Brain Sciences* 33:61–135.

Herne, Kaisa, Olli Lappalainen, and Elina Kestilae-Kekkonen. 2012. "Experimental comparison of direct, general, and indirect reciprocity."

mimeo. University of Turku. http://www.soc.utu.fi/laitokset/valtio-oppi/oppiaine/henkilokunta/herne_lappalainen_kestila.pdf (accessed January 22, 2013).

Herrmann, Benedikt and Henrik Orzen. 2008. "The appearance of homo rivalis: Social preferences and the nature of rent seeking." *Discussion Papers, The Centre for Decision Research and Experimental Economics, University of Nottingham.* 10.

Herrmann, Michael. 2012. "Voter uncertainty and failure of Duverger's law: an empirical analysis." *Public Choice* 151:63–90.

Herzberg, Roberta Q. and Rick K. Wilson. 1988. "Results on Sophisticated Voting in an Experimental Setting." *The Journal of Politics* 50(2):471–486.

Heukelom, Floris. 2007. "Kahneman and Tversky and the Origin of Behavioral Economics." *Tinbergen Institute Discussion Papers* 07-003/1:25.

Höffe, Otfried. 2005. "Zoon politikon." In *Aristoteles-Lexikon*. Kroener-Verlag, Stuttgart.

Hill, J. L., J. P. Reiter, and E. L. Zanutto. 2005. "A Comparison of Experimental and Observational Data Analyses." In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*, edited by A Gelman and X.-L. Meng, chapter 5. John Wiley & Sons, Ltd, Chichester, UK.

Hinich, Melvin J. 2008. "The Future of Analytical Politics." In *The Oxford Handbook of Political Economy*, edited by Barry R. Weingast and Donald Wittman, chapter 56. Oxford Handbooks Online.

Hinich, Melvin J. and Michael C. Munger. 1997. *Analytical politics*. Cambridge England ; New York, NY, USA: Cambridge University Press.

Hirschman, Albert O. 1964. "The Paternity of an Index." *The American Economic Review* 54:761–762.

Hodge, J. K. 2006. "Permutations of separable preference orders." *Discrete Applied Mathematics* 154:1478–1499.

Hodge, J. K. and P. Schwallier. 2006. "How does separability affect the desirability of referendum election outcomes?" *Theory and Decision* 61:251–276.

Hoechtl, Wolfgang, Rupert Sausgruber, and Jean-Robert Tyran. 2012. "Inequality aversion and voting on redistribution." *European Economic Review* 56:1406–1421.

Hofstadter, Douglas. 1985. *Metamagical Themas: Questing for the Essence of Mind and Pattern*. New York, Basic Books.

Holm, H. J. and P. Nystedt. 2010. "Collective Trust Behavior." *Scandinavian Journal of Economics* 112:25–53.

Holm, Hakan J., Sonja Opper, and Victor Nee. 2013. "Entrepreneurs Under Uncertainty: An Economic Experiment in China." *Management Science 01/2013* published online before print January 8, 2013.

Holt, C.A. and L. R. Anderson. 1999. "Agendas and strategic voting." *Southern Economic Journal* 65:622–629.

Holt, Charles and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92:1644–1655.

Holyoak, K.J. and D. Simon. 1999. "Bidirectional reasoning in decision making by constraint satisfaction." *Journal of Experimental Psychology* 128:3–31.

Holzinger, Katharina. 2009. "Vom ungeliebten Störenfried zum akzeptierten Paradigma? Zum Stand der (Neuen) Politischen Ökonomie in Deutschland." *Politische Vierteljahresschrift (PVS)* 50:539–576.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser. 2011. "The online laboratory: conducting experiments in a real labor market." *Experimental Economics* 14:399–425.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39:41–57.

Häring, Norbert and Olaf Storbeck. 2007. *Ökonomie 2.0: 99 überraschende Erkenntnisse*. Schäffer-Poeschel.

Hörl, Björn and Andreas Warntjen. 2005. "Built on quicksand? a decade of procedural spatial models on eu legislative decision-making." *Journal of European Public Policy* 12:592–606.

Hug, Simon. 2009. "Models of Multilateral Negotiations and Ratification." *mimeo. Paper prepared for presentation at the ISA conference March 1-5, 2005, Hawaii.* http://www.unige.ch/ses/spo/static/simonhug/momn.html (accessed February 26, 2012).

Huitt, Ralph K. 1954. "The Congressional Committee: A Case Study." *The American Political Science Review* 48:340–365.

Humphreys, Marcatan and Michael Laver. 2010. "Spatial Models, Cognitive Metrics, and Majority Rule Equilibria." *British Journal of Political Science* 40:11–30.

Humphreys, Paul. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method.* Oxford University Press.

Insko, Chester A., John Schopler, and Rick H. Hoyle. 1990. "Individual-Group Discontinuity as a Function of Fear and Greed." *Journal of Personality and Social Psychology* 58:68–79.

Iriberri, Nagore and Pedro Rey-Biel. 2011. "The role of role uncertainty in modified dictator games." *Experimental Economics* 14:160–180.

Isaac, R.M. and J.M. Walker. 1988a. "Communication and free-riding behavior: the voluntary contribution mechanism." *Economic Inquiry* 26:585–608.

Isaac, R.M. and J.M. Walker. 1988b. "Group size effects in public goods provision: the voluntary contributions mechanism." *Quarterly Journal of Economics* 103:179–199.

Isaac, R. Mark, James M. Walker, and Arlington W. Williams. 1994. "Group size and the voluntary provision of public goods : Experimental evidence utilizing large groups." *Journal of Public Economics* 54:1–36.

Jacquemet, Nicolas and Adam Zylbersztejn. 2011. "What drives failure to maximize payoffs in the lab ? A test of the inequality aversion hypothesis." Technical report, Documents de travail du Centre d'Economie de la Sorbonne 11036, Université Panthéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne.

Jeffrey, Milyo, David M. Primo, and Matthew L. Jacobsmeier. 2006. "Estimating the Impact of State Policies and Institutions with Mixed-Level Data." *Working Papers, Department of Economics, University of Missouri* 0603.

Jervis, R. 1997. *System Effects: Complexity in Political and Social Life*. Princeton: Princeton University Press.

Johnson, D. B. 1991. *Public choice: an introduction to the new political economy. .* Mountain View: Bristlecone Books.

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon. 2002. "Detecting failures of backward induction: Monitoring information search in sequential bargaining." *Journal of Economic Theory* 104:16–47.

Johnson, George. 2008. *The ten most beautiful experiments*. New York: Alfred A. Knopf.

Johnson, Robert. 2012. "Hedgefonds-Manager: "Ich konnte es nicht ertragen"." *Die ZEIT, Interview* 16.

Jones, Mark, Wonjae Hwang, and Juan Micozzi. 2009. "Government and Opposition in the Argentine Congress, 1989-2007: Understanding Inter-Party Dynamics through Roll Call Vote Analysis." *Journal of Politics in Latin America, German Institute of Global and Area Studies, Hamburg* 1:67–96.

Juettner, Julia, Benjamin Schulz, and Oliver Trenkamp. 2011. "Bundeswehr-Streichliste - Angst vor der Geisterstadt." *Spiegel - Magazin* 26.10.

Junge, Dirk. 2010. "Game Theoretic Models and the Empirical Analysis of EU Policy Making: Strategic Interaction, Collective Decisions, and Statistical Inference." In *Reform Processes and Policy Change: Veto Players and Decision-Making in Modern Democracies*, edited by Thomas König, George Tsebelis, and Marc Debus, pp. 165–186. New York.

Kagel, J.H., H. Sung, and E. Winter. 2010. "Veto power in committees: an experimental study." *Experimental Economics* 13:167–188.

Kagel, John H. and Alvin E. Roth. 1995. *The handbook of experimental economics*. Princeton, NJ: Princeton Univ. Press.

Kahneman, Daniel. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, Daniel, Paul Slovic, and Amos Tversky. 1982. *Judgment under uncertainty : heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–292.

Kaiser, Andre. 2007a. "James M. Buchanan, J. Gordon Tullock, The Calculus of Consent. Logical Foundations of Constitutional Democracy, Ann Arbor 1962." In *Schlüsselwerke der Politikwissenschaft*, edited by Steffen Kailitz, pp. 56–60. VS Verlag für Sozialwissenschaften, Wiesbaden.

Kaiser, A. 2007b. "Neuere Beitraege zur empirischen Institutionenanalyse: Von der Gegenstands-zur Problemorientierung." *Politische Vierteljahresschrift (PVS)* 48:119–135.

Kaiser, Andre, Jan Biela, and Annika Hennl. 2013. *Policy Making in Multilevel Systems*. ECPR Press.

Kaufmann, Stephan. 2012. "Appell an die Wirtschaftslehre: Lehre ohne akademische Prostitution." *Frankfurter Rundschau* printed edition April 4th.

Keeney, R.L. and H. Raiffa. 1976. *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Cambridge University Press.

Kehoe, C. and J. Pitkow. 1996. "Surveying the territory." *The World Wide Web Journal* 1:77–84.

Keiding, H. and B. Peleg. 2001. "Stable voting procedures for committees in economic environments." *Journal of Mathematical Economics* 36:117–140.

Kenning, P. and H. Plassmann. 2005. "NeuroEconomics: An overview from an economic perspective." *Brain Research Bulletin* 67:343–354.

Kesten, Yonit and Amir Pnueli. 1998. "Modularization and Abstraction: The Keys to Practical Formal Verification." *Mathematical Foundations of Computer Science* pp. 54–71.

Keynes, John Maynard. 1936. *The General Theory of Employmen*. Palgrave Macmillan.

Kinder, Donald R. and Thomas R. Palfrey. 1991. "An experimental political science? Yes, an experimental political science." *The Political Methodologist* 4:2–8.

King, Gary and Daniel Hopkins. 2007. *Extracting Systematic Social Science Meaning from Text*. Weatherhead Center for International Affairs, Harvard University.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

King, Gary and Margaret Roberts. 2012. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix." *mimeo. Harvard University.* gking.harvard.edu/files/robust.pdf (accessed March 3, 2013).

Kirchgässner, Gebhard. 2008. *Homo oeconomicus: Das okonomische Modell individuellen Verhaltens und seine Anwendung in den Wirtschafts- und Sozialwissenschaften*, volume 3. Tuebingen: Mohr Siebeck Gmbh & Co. K.

Kirchsteiger, Georg. 1994. "The role of envy in ultimatum games." *Journal of Economic Behavior and Organization* 25:373–389.

Kirsch, Guy. 2004. *Neue Politische Oekonomie*, volume 5. Stuttgart: Lucius und Lucius.

Kirstein, Roland and Dieter Schmidtchen. 2003. "Self-interest, Social Wealth, and Competition as a Discovery Procedure - A classroom experiment that makes the "invisible hand" visible." *Discussion Paper, Center for the Study of Law and Economics* 08.

Klein, Jonathan H. 1985. "The Abstraction of Reality for Games and Simulations." *The Journal of the Operational Research Society* 36:671–678.

Kleinbaum, David G., Lawrence L. Kupper, Azhar Nizam, and Keith E. Muller. 2007. *Applied Regression Analysis and Other Multivariable Methods*. Duxbury Press, 4th edition.

Knetsch, Jack L. and J. A. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics* 99:507–521.

König, Thomas. 2013. "Negative Preference Interdependence, Continuous Concessions and the Division of Reform Gains. Making the Lisbon Treaty Work." *mimeo. 6th Annual Conference on The Political Economy of International Organizations.* http://147.142.190.246/joomla/peio/files2013/papers/König%2030.09.2012.pdf (accessed March 23, 2013).

König, Thomas and Dirk Junge. 2009. "Why Don't Veto Players Use Their Power?" *European Union Politics* 10:507–534.

König, Thomas, Bernd Luig, and Stephan Marc Solomon. 2010. "Sachverstaendige und der Einfluss von Expertise auf Reformen: Eine raeumliche Analyse der Foederalismusreform II." *Perspektiven der Wirtschaftspolitik* 11:307–323.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. Houghton Mifflin.

Kocher, Martin G. and Matthias Sutter. 2005. "The Decision Maker Matters: Individual Versus Group Behaviour in Experimental Beauty-Contest Games." *Economic Journal, Royal Economic Society* 115:200–223.

Komorita, S. S. and C. D. Parks. 1995. "Interpersonal-Relations - Mixed-Motive Interaction." *Annual Review of Psychology* 46:183–207.

Koremenos, Barbara. 2008. "When, What, and Why do States Choose to Delegate?" *Law and Contemporary Problems* 71:151–192.

Kormendi, Roger C. and Charles R. Plott. 1982. "Committee Decisions under Alternative Procedural Rules - An Experimental Study Applying a New Nonmonetary Method of Preference Inducement." *Journal of Economic Behaviour and Organisation* 3:20.

Koukoumelis, A., M.V. Levati, and J. Weisser. 2012. "Leading by words: A voluntary contribution experiment with one-way communication." *Journal of Economic Behavior and Organization* 81:379–390.

Kramer, Gerald H. 1972. "Sophisticated voting over multidimensional choice spaces." *The Journal of Mathematical Sociology* 2:165–180.

Kranich, L., A. Perea, and H. Peters. 2005. "Core Concepts for Dynamic TU Games." *International Game Theory Review* 7:43–61.

Krantz, J., J. Ballard, and J. Scher. 1997. "Comparing the results of laboratory andworld-wide web samples on the determinants of female attractiveness." *Behavior Research Methods, Instruments, and Computers* 29:264–269.

Krehbiel, Keith. 1986. "Sophisticated and Myopic Behavior in Legislative Committees: An Experimental Study." *American Journal of Political Science* 30(3):542–561.

Krehbiel, Keith. 1988. "Spatial Models of Legislative Choice." *Legislative Studies Quarterly* 13:259–319.

Krehbiel, Keith, Adam Meirowitz, and Jonathan Woon. 2004. "Testing Theories of Law-making." *Research Papers, Stanford University* 1860.

Kriss, Peter H. and Roberto A. Weber. 2012. "Economic Laboratory Experiments on Organizational Formation and Change." In *Handbook of Economic Organization: Intergrating Economic and Organization Theory*, edited by Anna Grandori. Edward Elgar.

Kröning, Daniel and Ofer Strichman. 2008. *Decision procedures : an algorithmic point of view*. Berlin: Springer.

Krugman, Paul. 2010. "Nobelpreisträger Krugman will alte VWL-Weisheiten "beerdigen"." In *Interview mit dem Handelsblatt, Ausgabe vom 11.01.*

Kugler, Tamar, Gary Bornstein, Matthias Sutter, and Martin G. Kocher. 2007. "Trust between individuals and groups: Groups are less rusting than individuals but just as trustworthy." *Journal of Economic Psychology* 28:646–657.

Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers. 1997. "Racial Prejudice and Attitudes Toward Affirmative Action." *American Journal of Political Science* 41:402–419.

Kurzban, R. and D. Houser. 2005. "Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations." *Proceedings of the National Academy of Sciences of the United States of America* 102:1803–1807.

Kuziemko, Ilyana, Ryan W. Buell, Taly Reich, and Michael I. Norton. 2011. “"Last-place Aversion": Evidence and Redistributive Implications.” *NBER Working Paper Series* 17234. National Bureau of Economic Research, Inc.

Laband, David and Richard Biel. 1999. “Are Economists More Selfish than Other 'Social' Scientists?” *Public Choice* 100:1–2, 85–101.

Lacy, Dean. 2001a. “Nonseparable Preferences, Measurement Error, and Unstable Survey Responses.” *Political Analysis* 9:1–21.

Lacy, Dean. 2001b. “A theory of nonseparable preferences in survey responses.” *American Journal of Political Science* 45:239–258.

Lacy, Dean. 2003. “Dependent preferences and over-time instability in survey responses.” *Presented at the annual meeting of the American political science association* .

Lacy, D. and Niou. 1998a. “Nonseparable Preferences, Issue Linkage, and Economic Sanctions.” *mimeo. Presented at the Annual Meeting of the American Political Science Association, Boston, MA, September 3-6.* citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.8.8872&rep=rep1&type=pdf (accessed December 18, 2012).

Lacy, Dean and Emerson M. S. Niou. 1994. “Preferences and a Problem with Referendums.” *mimeo. Presented at the Annual Meeting of the Public Choice Society and Economic Science Association, April 8-10, Austin, Texas.*

Lacy, D. and E. M. S. Niou. 1998b. “Elections in double-member districts with nonseparable voter preferences.” *Journal of Theoretical Politics* 10:89–110.

Lacy, Dean and Emerson M. S. Niou. 2000. “A problem with referendums.” *Journal of Theoretical Politics* 12:5–31.

Lai, Ernest K. and Wooyoung Lim. 2012. “Authority and communication in the laboratory.” *Games and Economic Behavior* 74:541–560.

Lalive, Rafael and Armin Schmutzler. 2008. “Exploring the Effects of Competition for Railway Markets.” *International Journal of Industrial Organisation* 26:443–458.

Lane, Jan-Erik. 2000. “The New Institutionalism.” In *The Public Sector: Concepts, Models and Approaches*, chapter 10. SAGE Publications, London.

Lane, Jan-Erik and Svante Ersson. 2000. *The New Institutional Politics: Performance and Outcomes*. London: Routledge.

Lang, Jerome and Lirong Xia. 2009. “Sequential composition of voting rules in multi-issue domains.” *Mathematical Social Sciences* 57:304–324.

Larsson, Rikard. 1993. “The Handshake between Invisible and Visible Hands: Toward a Tripolar Institutional Framework.” *International Studies of Management and Organization* 23:87–106.

Lask, Reinhard. 2010. "Anbau-Gegner sehen die "Stadthalle im Chaos" versinken." *Rgein-Neckar-Zeitung* 15.06.

Lave, Charles and James G. March. 1975. *An Introduction to Models in the Social*. New York: Harper and Row.

Laver, M., K. Benoit, and J. Garry. 2003. "Extracting policy positions from political texts using words as data." *American Political Science Review* 97:311–331.

Laver, M. and M. Schilperoord. 2007. "Spatial models of political competition with endogenous political parties." *Philosophical Transactions of the Royal Society B-Biological Sciences* 362:1711–1721.

Laver, Michael and Norman Schofield. 1998. *Multiparty government : the politics of coalition in Europe*. Ann Arbor: University of Michigan Press.

Le Breton, M. and A. Sen. 1999. "Separable preferences, strategyproofness, and decomposability." *Econometrica* 67:605–628.

Leblebici, Huseyin and Gerald R. Salancik. 1981. "Effects of Environmental Uncertainty on Information and Decision Processes in Banks." *Administrative Science Quarterly* 26:578–596.

Ledyard, J. 1995. "Public goods: A survey of experimental research." In *The Handbook of Experimental Economics*, edited by J. H. Kagel and A.E. Roth, pp. 111–194. NJ: Princeton University Press.

Lee, Jinkwon. 2008. "The Effect of the Background Risk in a Simple Chance Improving Decision Model." *The Journal of Risk and Uncertainty* 36:19–41.

Lehrer, Ehud and Marco Scarsini. 2011. "On the Core of Dynamic Cooperative Games." *Cornell University Library* arXiv:1203.2832.

Lerner, A. P. and H. W. Singer. 1937. "Some Notes on Duopoly and Spatial Competition." *The Journal of Political Economy, University of Chicago Press* 45:145–186.

Lesk, A. M. 2000. "The unreasonable effectiveness of mathematics in molecular biology." *The Mathematical Intelligencer* 22:28–37.

Leuffen, Dirk, Susumu Shikano, and Stefanie Walter. 2012. "Measurement and Data Aggregation in Small-N Social Scientific Research." *European Political Science* 2012/03/30/online.

Levin, Simon A. 2009. *Games, groups, and the global good*. Berlin ; New York: Springer.

Levine, David K. 1997. "Modeling Altruism and Spitefulness in Experiments." *Journal of Economic Literature* .

Levine, John M. and Richard L. Moreland. 1998. "Small Groups." In *The Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, volume II, chapter Part IV, 26, pp. 415–469. Boston: McGraw-Hill, 4th edition.

Levitt, Steven D. and Stephen J. Dubner. 2009. *Freakonomics*. Harper.

Levitt, Steven D. and Stephen J. Dubner. 2011. *SuperFreakonomics: Global Cooling, Patriotic Prostitutes, and Why Suicide Bombers Should Buy Life Insurance*. William Morrow Paperbacks.

Levitt, Steven D. and John A. List. 2006. "What Do Laboratory Experiments Tell Us About the Real World?" *mimeo. University of Chicago and NBER.* http://pricetheory.uchicago.edu/levitt/Papers/jep%20revision%20Levitt%20&%20 List.pdf (accessed November 6, 2010).

Levitt, Steven D. and John A. List. 2007. "What do laboratory experiments measuring social preferences reveal about the real world?" *Journal of Economic Perspectives* 21:153–174.

Levy, Haim. 2006. *Stochastic Dominance: Investment Decision Making under Uncertainty, Springer*. Springer US.

Lewin, K. 1936. *Principles of topological psychology*. New York: McGraw-Hill.

Lichbach, Mark I. 2001. *Is Rational Choice Theory All of Social Science? .* Ann Arbor: University of Michigan Press.

Lijphart, Arend. 1971. "Comparative politics and the comparative method." *American Political Science Review* 65:682–693.

Lijphart, A. 1999. *Patterns of Democracy*. New Haven: Yale University Press.

Lind, Edgar Allen. 1988. *The social psychology of procedural justice*. New York: Plenum Press.

Lindgren, Kristian. 1991. "Evolutionary Phenomenon in Simple Dynamics." In *Artificial Life II*, edited by Christopher G. Langton, Charles Taylor, J. Doyne Farmer, and Steen Rasmussen. Addison-Wesley, Reading, MA.

Lipset, Seymour M. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53:69–105.

List, John A. 2009. "Social Preferences: Some Thoughts from the Field." *Annual Review of Economics* pp. 563–579.

Lockhead, Gregory and James Pomerantz. 1991. *The Perception of Structure [Paperback]*. American Psychological Association.

Lomborg, B. 1996. "Nucleus and shield: The evolution of social structure in the iterated Prisoners dilemma." *American Sociological Review* 61:278–307.

Long, Ngo Van. 2010. *A survey of dynamic games in economics*. Singapore ; London: World Scientific.

Long, Roderick T. 2006. "Realism and abstraction in economics: Aristotle and Mises versus Friedman." *The Quarterly Journal of Austrian Economics* 9:3–23.

Lopomo, Giuseppe and Efe Ok. 2001. "Bargening, Interdependence and the Rationality of fair division." *Levine's Bibliography, UCLA Department of Economics* .

Lowell, A. Lawrence. 1910. "The Physiology of Politics." *American Political Science Review* 4:1–15.

Loxley, A. 1997. *Collaboration in Health and Welfare*. Jessica Kingsley Publishers, London.

Luce, D. 1959. *Individual Choice Behavior*. New York: Wiley.

Lucking-Reiley, David H. 1999. "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet." *American Economic Review* 89:1063–1080.

Löwenstein, G., L. Thompson, and M. Bazerman. 1989. "Social utility and decision making in interpersonal contexts." *Journal of Personality and Social Psychology* 53:426–441.

MacKenzie, Scott. 2012. "Within-subjects vs. Between- subjects Designs: Which to Use?" *mimeo. Dept. of Computer Science, York University.* http://www.yorku.ca/mack/RN-Counterbalancing.html (accessed November 28, 2012).

Maier-Rigaud, Frank P. and Jose Apesteguia. 2003. "The Role of Choice in Social Dilemma Experiments." *mimeo. Economic Discussion Papers, University of Bonn.* bgse22. http://www.wiwi.uni-bonn.de/bgsepapers/bonedp/bgse22_2003.pdf (accessed December 11, 2011).

Malecki, Michael and Matthew Gabel. 2007. "A Survey Experiment on Satisfaction with Democracy: How question order shapes respondents' evaluations of satisfaction with democracy in the EU." *mimeo. Paper presented at the annual meeting of the American Political Science Association, Chicago.* http://www.allacademic.com/meta/p210366_index.html (accessed March 13, 2012).

Manhart, Klaus. 2011. "IDC-Studie zum Datenwachstum - Doppeltes Datenvolumen alle zwei Jahre." *CIO* July.

Mankiw, N. G. and Mark Taylor. 2010. *Economics: Special Edition With Global Economic Watch*. Southwestern Cengage Learning.

Mann, H. B. and D. R. Whitney. 1947. "On a test whether one of two random variables is stochastically larger than the other." *Annals of Mathematical Statistics* 18:50–60.

Mann, Thomas. 2005. "REACH-Kompromiss nachbessern!" *Pressemeldung* 17th November.

Mannes, Albert E. 2009. "Are We Wise About the Wisdom of Crowds? The Use of Group Judgments in Belief Revision." *Management Science* 55:1267–1279.

March, James G. and Johan P. Olsen. 2006. "The Logic of Appropriateness." In *The Oxford Handbook of Public Policy*, edited by Michael Moran, Martin Rein, and Robert E. Goodin, pp. 689–708. Oxford: Oxford University Press.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953 to 1999." *Political Analysis* 10:134–153.

Martin, D. W. 2008. *Doing psychology experiments*, volume 7. Pacific Grove, CA: Brooks/Cole.

Marwell, Gerald and Ruth E. Ames. 1981. "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods, IV." *Journal of Public Economics* 15:295–310.

Mason, W. and S. Suri. 2012. "Conducting behavioral research on Amazon's Mechanical Turk." *Behavior Research Methods* 44:1–23.

Matis, Herbert. 2007. "Das Menschenbild der klassischen Nationalökonomie." In *Das Menschenbild in der Ökonomie: eine verschwiegene Voraussetzung*, edited by Reinhard Pirker, pp. 105–140. Wien: Löcker.

Mattila, Mikko. 2004. "Contested decisions: Empirical analysis of voting in the European Union Council of Ministers." *European Journal of Political Research* 43:29–50.

May, Kenneth O. 1952. "A set of independent necessary and sufficient conditions for simple majority decisions." *Econometrica* 20:680–684.

Mayring, Philipp. 2002. *Einfuehrung in die qualitative Sozialforschung : eine Anleitung zu qualitativem Denken*. Weinheim ; Basel: Beltz.

Mayring, Philipp. 2010. *Qualitative Inhaltsanalyse : Grundlagen und Techniken*. Weinheim ; Basel: Beltz.

McCarty, Nolan M. and Adam Meirowitz. 2007. *Political game theory : an introduction*. Cambridge [u.a.]: Cambridge University Press.

McConnachie, Alistair. 2000. "Direct democracy - Government of the People, by the People, for the People." *Sovereignty* March issue.

McCubbins, M., R. G. Noll, and B. R. Weingast. 1987. "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics and Organization* .

McCubbins, Mathew, Roger G. Noll, and Barry R. Weingast. 1989. "Structure and Process; Politics and Policy: Administrative Arrangements and the Political Control of Agencies." *Virginia Law Review* .

McCubbins, Mathew and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols vs. Fire Alarms." *American Journal of Political Science* .

McDermott, Rose. 2002. "Experimental Methology in Political Science." *Political Analysis* 10:4:325–342.

McElduff, Fiona, Mario Cortina-Borja, Shun-Kai Chan, and Angie Wade. 2010. "When t-tests or Wilcoxon-Mann-Whitney tests won't do." *Advances in Physiology Education* 34:128–133.

McFadden, Daniel. 2006. "Free Markets and Fettered Consumers." *American Economic Review, American Economic Association* 96:5–29.

McFadden, Daniel. 2013. "The New Science of Pleasure." *NBER Working Paper* 18687.

McFadden, Daniel L. 1974. "Conditional logit analysis of qualitative choice behavior." In *Frontiers in econometrics*, edited by P. Zarembka, pp. 105–142. New York: Academic Press.

McFadden, Daniel L. 1976. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement* 5:363–390.

McFadden, Daniel L. 1982. "Econometric Models of Probabilistic Choice." In *Structural analysis of discrete data with econometric applications*, edited by C. Manski and D. L. McFadden, pp. 198–272. Cambridge, MIT Press.

McKelvey, Richard D. 1976. "Intransitivities in multidimensional voting models and some implications for agenda control." *Journal of Economic Theory* 12:472–482.

McKelvey, Richard D. and Peter C. Ordeshook. 1981. "Experiments on the Core: Some Disconcerting Results for Majority Rule Voting Games." *Journal of Conflict Resolution* 25:473–482.

McKelvey, Richard D. and Peter C. Ordeshook. 1982. "Two-Candidate Elections Without Majority Rule Equilibria: An Experimental Study." *Simulation and Games* 13:311–335.

McKelvey, Richard D. and Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10:6–38.

McKelvey, Richard D. and Thomas R. Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *Japanese Economic Review* 47:186–209.

McKelvey, Richard D. and Thomas R. Palfrey. 1998. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics* 1:9–41.

McNulty, Roy. 2011. "Realising the Potential of GB Rail." *Report of the Rail Value for Money Study* .

Merkel, Wolfgang. 2010. "Volksentscheide - wer ist das Volk." *Neue Gesellschaft Frankfurter Hefte* 12.

Merrill, Samuel and Bernard Grofman. 1999. *A Unified Theory of Voting: Directional and Proximity Spatial Models*. Cambridge University Press.

Mertins, Vanessa. 2008. "The Effects of Procedures on Social Interaction: A Literature Review." *Discussion Paper Series, Institute for Labour Law and Industrial Relations in the European Community, University Trier* 06.

Mickan, Sharon and Sylvia Rodger. 2000. "Characteristics of effective teams: a literature review." *Australian Health Review* 23.

Miljkovic, Sladjana Lj. 2012. *Iterative methods for computing generalized inverses of matrices*. Ph.D. thesis, University of Nis, Faculty of Natural Sciences and Mathematics.

Miller, Gary J. 1997. "The Impact of Economics on Contemporary Political Science." *Journal of Economic Literature* XXXV:1173–1204.

Miller, G. J. and T. H. Hammond. 1990. "Committees and the Core of the Constitution." *Public Choice* 66:201–227.

Miller, G. J., T. H. Hammond, and C. Kile. 1996. "Bicameralism and the core: An experimental test." *Legislative Studies Quarterly* 21:83–103.

Miller, Gary J. and Joe A. Oppenheimer. 1982. "Universalism in Experimental Committees." *The American Political Science Review* 76(3):561–574.

Miller, Luis and Christoph Vanberg. 2014. "Decision costs in legislative bargaining: An experimental analysis." *Public choice* (accepted). alt: 2011. Discussion Papers, University of Oxford, Nuffield College. 2011002.

Milyo, Jeffrey. 2000a. "Logical deficiencies in spatial models: A constructive critique." *Public Choice* 105:273–289.

Milyo, Jeffrey. 2000b. "A problem with Euclidean preferences in spatial models of politics." *Economics Letters* 66:179–182.

Mitchell, Mark L. 2003. *Research Design Explained*. Thomson Learning.

Mitchell, Terence R. 1985. "Review and Evaluation of Research Studies on Negotiation Behaviour." *Journal of Occupational Behaviour* 6:85–91.

Mitchell, William C. 1999. "Political science and public choice: 1950-70." *Public Choice* 98:237–249.

Moe, Terry. 1990. "Political Institutions - The Neglected Side of the Story." *Journal of Law, Economics and Organizatio* 6:39–77.

Monroe, Kristen R. 2005. *Perestroika! The Raucous Rebellion in Political Science*. New Haven-London: Yale University Press.

Morgan, T. Clifton. 1984. "A Spatial Model of Crisis Bargaining." *International Studies Quarterly* 28:407–426.

Morgan, T. C. 1990. "Issue linkage in international crisis bargaining." *American Journal of Political Science* 34:311–333.

Morgan, T. C. 1994. *Untying the Knot of War*. Ann Arbor, MI: University of Michigan Press.

Morris, Dick. 1999. *The New Prince*. Los Angeles: Renaissance Books.

Morris, Stephen and Hyun Song Shin. 2002. "Measuring Strategic Uncertainty." *mimeo. Working paper, Yale University.* http://www.princeton.edu/˜hsshin/www/barcelona.pdf (accessed March 3, 2012).

Morton, Rebecca and Kenneth C. Williams. 2006. "Experimentation in Political Science." In *Cambridge Handbook of Experimental Political Science*, edited by Janet Box-Steffensmeier, David Collier, and Henry Brady, pp. 1–27. Oxford: Oxford University Press.

Morton, Rebecca B. 1999. *Methods and models : a guide to the empirical analysis of formal models in political science*. Cambridge: Cambridge University Press.

Morton, Rebecca B. and Kenneth C. Williams. 1999. "Information Asymmetries and Simultaneous Versus Sequential Voting." *American political science Review* 93:51–67.

Morton, Rebecca B. and Kenneth C. Williams. 2010. *Experimental political science and the study of causality : from nature to the lab*. New York: Cambridge University Press.

Morton, Rebecca B. and Kenneth C Williams. 2011. "Electoral Systems and strategic voting (laboratory election experiments)." In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, pp. 369–383. New York: Cambridge Unviersity Press.

Morton, Rebecca B. and Kenneth C. Williams. 2012. "Experimente in der Politischen Ökonomie." In *Jahrbuch für Handlungs- und Entscheidungstheorie*, edited by Thomas Bräuninger, Bächtiger, and Susumu Shikano, number Band 7: Experiment und Simulation, chapter 1, pp. 81–104. VS Verlag für Sozialwissenschaften, Wiesbaden.

Moser, P. 1997. "A theory of the conditional influence of the European Parliament in the cooperation procedure." *Public Choice* 91:333–350.

Moses, Jonathon Wayne and Torbjoern L. Knutsen. 2007. *Ways of knowing : competing methodologies and methods in social and political research*. New York: Palgrave Macmillan.

Motulsky, Harvey J. 2010. *Intuitive Biostatistics*. Oxford University Press, 2nd edition.

Möslinger, Astrid. 2010. "Heidelberg: Klotzen für die Ewigkeit." *Meier - Magazin* 07 - 10:32–33.

Mueller, D. 2003. *Public Choice III*. Cambridge University Press, New York.

Muren, Astri and Roger Pyddoke. 1999. "Does Collusion without Communication Exist?" *Research Papers in Economics, Stockholm University, Department of Economics.* 11.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment and Decision Making* 6:771–781.

Nagel, R., A. Bosch, J. Montalvo, and A. Satorra. 2002. "One, two, [three], infinity: newspaper and lab beauty-contest experiments." *American Economic Review* 92:1687–1701.

Nash, John F. jr. 1950. "The Bargaining Problem." *Econometrica* 18:155–162.

Neale, M. A. and M. Bazerman. 1991. *Cognition and rationality in negotiation*. New York: Free Press.

Newson, Roger. 2002. “Parameters behind "nonparametric" statistics: Kendall’s tau, Somers’ D and median differences.” *Stata Journal* 2:45–64.

Newson, Roger. 2006a. “Confidence intervals for rank statistics: Somers’ D and extensions.” *The Stata Journal* 6:309–334.

Newson, Roger. 2006b. “Efficient Calculation of Jackknife Confidence Intervals for Rank Statistics.” *Journal of Statistical Software* 15.

Newson, Roger. 2013. “Interpretation of Somers’ D under four simple models.” *mimeo. King’s College London.* www.imperial.ac.uk/nhli/r.newson/miscdocs/intsomd1.pdf (accessed April 22, 2013).

Nice. 2001. “Treaty of Nice - Amending the treaty on European union, the treaties establishing the Eurpean communities and certain related acts.” *Official Journal of the European Communities* 2001/C80/01. signed on 26th February.

Nicklisch, Andreas and Irenaeus Wolff. 2012. “On the Nature of Reciprocity: Evidence from the Ultimatum Reciprocity Measure.” *Research Paper Series, Thurgauer Wirtschaftsinstitut, Universitaet Konstanz.* 65.

Niemi, Richard G. and Arthur Q. Frank. 1985. “Sophisticated voting under the plurality procedure: A test of a new definition.” *Journal of Theory and Decision* 19:151–162.

Niskanen, William A. 1971. *Bureaucracy and Representative Government*. Chicago: Aldine.

Noel, Hans. 2010. “Ten Things Political Scientists Know that You Don’t.” *Political science and practical Politics Forum, Georgetown University* 8. http://www9.georgetown.edu/faculty/hcn4/Downloads/Noel_Forum.pdf (accessed July 7, 2011).

North, Douglass. 2008. “Violence.” *Keynote Address, Upton Forum* www.beloit.edu/upton/assets/North.pgs.vol.I.pdf (accessed October 22, 2011).

North, Douglass Cecil. 1990. *Institutions, institutional change, and economic performance*. Cambridge ; New York: Cambridge University Press.

North, Douglass C. 1997. “Nobel Prize Lecture: Economic Performance through Time.” In *From Nobel Lectures, Economics 1991-1995*, edited by Torsten Persson. Singapore: World Scientific Publishing Co.

North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.

Nosofsky, Robert. 1985. “Overall similarity and the identification of separable-dimension stimuli: A choice model analysis.” *Perception and Psychophysics* 38:415–432.

Nullmeier, Frank and Tanja Pritzlaff. 2009. “The implicit normativity of political practices. Analyzing the dynamics and power relations of committee decision-making.” *Critical Policy Studies* 3:357–374.

Nutzinger, Hans G. 2010. “Nobelpreis in Wirtschaftswissenschaften für Elinor Ostrom: Ein Überblick über ihr ökonomisches Hauptwerk.” *MAGKS Discussion Paper Series in Economics* 24.

Oaksford, M. and N. Chater. 1992. “Bounded rationality in taking risks and drawing inferences.” *Theory and Psychology* 2:225–230.

Oaksford, M. and N. Chater. 1998. *Rationality in an uncertain world: Essays on the cognitive science of human reasoning*. Psychology Press.

Oates, Wallace E. 2005. “Toward A Second-Generation Theory of Fiscal Federalism.” *International Tax and Public Finance* 12:349–373.

Ockenfels, Axel. 1999. *Fairness, Reziprozität und Eigennutz : Oekonomische Theorie und experimentelle Evidenz*, volume 108. Tuebingen: Mohr Siebeck.

Ockenfels, Axel. 2007. “Daten statt Dogmen.” In *Ökonomie 2.0: 99 überraschende Erkenntnisse*. Schäffer-Poeschel.

Offerman, T. 2002. “Hurting hurts more than helping helps.” *European Economic Review* 46:1423 1437.

Offerman, T., J. Sonnemans, and A. Schram. 1996. “Value orientations, expectations, and voluntary contributions in public goods.” *Economic Journal* 106:817–845.

Ok, Efe A. and Levent Kockesen. 2000. “Negatively interdependent preferences.” *Social choice and welfare* 17:533–558.

Olken, Benjamin A. 2008. “Direct democracy and local public goods evidence from a field experiment in Indonesia.” *NBER Working Paper Series* 14123.

Olsen, Johan P. 2007. “Organization theory, public administration, democratic governance.” *ARENA Working Papers* 1.

Olson, Mancur. 1965. *The logic of collective action; public goods and the theory of groups*, volume 124. Cambridge, Mass.: Harvard University Press.

Olum, Yasim. 2004. “Modern management theories and practices: an critical overview.” presented at the 15th East African Central Banking Course at Kenya School of Monetary Studies. http://unpan1.un.org/intradoc/groups/public/documents/AAPAM/UNPAN025765.pdf (accessed November 25, 2013).

Opp, Karl-Dieter. 1999. “Contending Conceptions of the Theory of Rational Action.” *Journal of Theoretical Politics* 11:171–202.

Opp, Karl-Dieter. 2004. “Die Theorie rationalen Handelns im Vergleich mit alternativen Theorien.” In *Paradigmen der akteurszentrierten Soziologie*, edited by Manfred Gabriel, pp. 43–68. VS Verlag für Sozialwissenschaften, Wiesbaden.

Ordeshook, Peter C. 1986. *Game theory and political theory : an introduction*. Cambridge Cambridgeshire ; New York: Cambridge University Press.

Ordeshook, Peter C. and Thomas R. Palfrey. 1988. "Agendas, Strategic Voting, and Signaling with Incomplete Information." *American Journal of Political Science* 32:441–466.

Ordeshook, Peter C. and Mark Winer. 1980. "Coalitions and Spatial Policy Outcomes in Parliamentary Systems: Some Experimental Results." *American Journal of Political Science* 24(4):730–752.

Ortlieb, Claus Peter. 2010. "Interview: Ökonomie ist eigentlich keine Wissenschaft." *FAZ - Frankfurter Allgemeine Zeitung* Ausgabe vom 08.05.

Ortman, Andreas and Gerd Gigerenzer. 2000. "Reasoning in Economics and Psychology: Why Social Context Matters." In *Cognition, Rationality, and Institutions*, edited by M.E. Streit, U. Mummert, and D. Kiwit, pp. 131–145. Springer.

Ortmann, A. 2003. "Charles R. Plott's collected papers on the experimental foundations of economic and political science." *Journal of Economic Psychology* 24:555–575.

Ortmann, Andreas. 2005. "Field experiments in Economics: some methodological caveats." In *Field Experiments in Economics*, edited by Glenn W. Harrison, Jeffrey Carpenter, and John A. List, volume 10 of *Research in Experimental Economics*, pp. 51–70. Emerald Group Publishing Limited.

Ostrom, Elinor. 1986. "An agenda for the study of institutions." *Public Choice* 48:3–25.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity*. Princeton: Princeton University Press.

Ostrom, Elinor. 2006. "Converting Threats into Opportunities." *Political Science and Politics* 39:3–12.

Ostrom, Elinor. 2008. "Polycentric Systems as One Approach for Solving Collective-Action Problems." *Social Science Research Network* available at SSRN: http://ssrn.com/abstract=1304697 (accessed February 1, 2013).

Ostrom, Vincent. 1999. "Polycentricity." In *Polycentricity and Local Public Economies: Readings from the Workshop in Political Theory and Policy Analysis*, edited by Michael McGinnis, pp. 52–74. Ann Arbor, MI: University of Michigan Press.

Oviedo, J. 2000. "The core of a repeated n-person cooperative game." *European Journal of Operational Research* 127:519–524.

Palacios-Huerta, Ignacio and Oscar Volij. 2008. "Experientia docet: professionals play minimax in laboratory experiments." *Econometrica* 76:71–115.

Palfrey, T. and J. E. Prisbrey. 1997. "Anomalous Behavior in Public Goods Experiments: How Much and Why?" *American Economic Review* 87:829–846.

Palfrey, Thomas R. 1989. "A Mathematical Proof of Duverger's Law." In *Models of Strategic Choice in Politics*, edited by Peter C. Ordeshook, pp. 69–92. Ann Arbor: University of Michigan Press.

Palfrey, Thomas R. 2009. "Laboratory Experiments in Political Economy." *Annual Review of Political Science* 12:379–388.

Paolacci, Gabriele. 2012. "Experimental Turk." *A blog on social science experiments using Amazon Mechanical Turk* experimentalturk.wordpress.com.

Paolacci, G., J. Chandler, and P. G. Ipeirotis. 2010. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making* 5:411–419.

Pearl, J. 2009. "Causal inference in statistics: An overview." *Statistics Surveys* 3:96–146.

Pedersen, Anne Reff, Karina Sehested, and Eva Sörensen. 2011. "Emerging Theoretical Understanding of Pluricentric Coordination in Public Governance." *The American Review of Public Administration* 41:375–394.

Peleg, B. and P. Sudhoelter. 2003. *Introduction to the Theory of Cooperative Games*. Boston: Kluwer Academic.

Persky, Joseph. 1995. "The Ethology of Homo Economicus." *Journal of Economic Perspectives, American Economic Association* 9:221–231.

Persson, Torsten and Guido Enrico Tabellini. 1992. *Federal fiscal constitutions. Pt.1, Risk sharing and moral hazard*. London: Centre for Economic Policy Research.

Perttunen, C.D., D.R. Jones, and B.E. Stuckman. 1993. "Lipschitzian optimization without the Lipschitz constant." *Journal of Optimization Theory and Applications* 79:175–181.

Peters, B. G. 2000. *Institutional theory in political science*. London: Continuum. repr. ed.

Pich, Michael T., Christoph H. Lock, and Arnoud de Meyer. 2002. "On Uncertainty, Ambiguity, and Complexity in Project Management." *Management Science* 48:10008–1023.

Pigliucci, Massimo. 2013. "Rationally Speaking." *A blog on the Philosophy of Science* http://rationallyspeaking.blogspot.de (accessed October 23, 2012).

Plott, C.R. 1979. "The application of laboratory experimental methods to public choice." In *Collective decision making: Applications from public choice theory*, edited by C.S. Russell, pp. 137–160. Baltimore: Johns Hopkins University Press.

Plott, Charles R. 1967. "A Notion of Equilibrium and Its Possiblity Under Majority Rule." *American Economic Review* 57:787–806.

Plott, C. R. 1991. "Will Economics Become an Experimental Science." *Southern Economic Journal* 57:901–919.

Plott, Charles R. and Vernon L. Smith. 2008. *Handbook of experimental economics results. [Volume 1]*. Amsterdam ; New York: North Holland.

Plott, C. R. and K. Zeiler. 2011. "The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Reply." *American Economic Review* 101:1012–1028.

Poisson, S. D. 1837. *Recherches sur la probabilité des jugements en matière criminelle et matière civile*. Paris: Bachelier.

Prasnikar, Vesna. 1997. "Learning the Decision Rules in Ultimatum Games." *Working paper* .

Pratt, J. W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* XXXII:122–36.

Pritzlaff, Tanja. 2006. *Entscheiden als Handeln*. Campus Verlag, Frankfurt.

Proksch, Sven-Oliver and Jonathan B. Slapin. 2009. "WORDFISH: Scaling Software for Estimating Political Positions from Texts." *http://www.wordfish.org*. Version 1.3 (22 January 2009).

Provis, Chris. 2000. "Ethics, Deception and Labor Negotiation." *Journal of Business Ethics* 28:145–158.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42:427–460.

Quattrone, George A. and Amos Tversky. 1988. "Contrasting Rational and Psychological Analyses of Political Choice." *The American Political Science Review* 82(3):719–736.

Rabin, M. 1993. "Incorporating Fairness into Game-Theory and Economics." *American Economic Review* 83:1281–1302.

Rabin, M. 1998. "Psychology and economics." *Journal of Economic Literature* 36:11 46.

Radke, Sabine. 2011. *Verkehr in Zahlen*. DVV Media Group GmbH, Hamburg. 40. Jahrgang.

Ragin, C. C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, Charles C. 2006. "Set Relations in Social Research: Evaluationg Their Consistency and Coverage." *Political Analysis* 14:291–310.

Ragusa, Jordan, Joshua Huder Huder, and Nate Birkhead. 2012. "Rule 22." In *Dedicated to Unlimited Debate about American Politics*.

Rand, D. G. 2012. "The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments." *Journal of Theoretical Biology* 299:172–179.

Rawls, John. 1971. *A Theory of Justice*. Belknap. Original 1971, überarbeitete dt. Fassung 1975.

Ray, Debraj. 2007. *A game-theoretic perspective on coalition formation*. Oxford ; New York: Oxford University Press.

Redner, R. and H. Walker. 1984. "Mixture Densities, Maximum Likelihood and the EM Algorithm." *Siam Review* 26:195–239.

Reid, Joseph D. Jr. 1987. "The Theory of Sharecropping: Occams Razor and Economic Analysis Winter 1987 19(4):." *History of Political Economy* 19:551–569.

Reips, U. 2000. "The web experiment method: advantages, disadvantages, and solutions." In *Psychological Experiments on the Internet*, edited by M. Birnbaum, pp. 89–120. Academic Press, San Diego.

Resnick, P., K. Kuwabara, R. Zeckhauser, and E. Friedman. 2000. "Reputation Systems." *Communications of the ACM* 43:45–48.

Richards, D. 1994. "Intransitivities in Multidimensional Spatial Voting - Period three Implies Chaos." *Social Choice and Welfare* 11:109–119.

Riesch, Hauke. 2010. "Simple or simplistic? Scientists' views on Occam's razor." *Theoria* 25:75–90.

Riker, William H. 1980. "Implications from the Disequilibrium of Majority Rule for the Study of Institutions." *The American Political Science Review* 74:432–446.

Riker, William H. 1982a. *Liberalism Against Populism*. San Francisco: W.H. Freeman.

Riker, William H. 1982b. "The Two-Party System and Duverger's Law: An Essay on the History of Political Science." *American Political Science Review* 76:753–766.

Risse, Mathias. 2004. "Arguing for Majority Rule." *The Journal of Political Philosophy* 12:41–64.

Rosch, Eleanor and Barbara B. Lloyd. 1978. *Cognition and categorization*. Hillsdale, NJ: John Wiley & Sons Inc.

Rosenthal, Robert and Ralph L. Rosnow. 2009. *Artifacts in behavioral research: Robert Rosenthal and Ralph L. Rosnow's classic books : a re-issue of Artifact in behavioral research, Experimenter effects in behavioral research and the volunteer subject*. New York: Oxford University Press.

Roth, A. E. and I. Erev. 1995. "Learning in Extensive-Form Games - Experimental-Data and Simple Dynamic-Models in the Intermediate-Term." *Games and Economic Behavior* 8:164–212.

Royston, P.J. 1992. "Approximating the Shapiro-Wilk W-Test for Non-normality." *Statistics and Computing* 2:117–119.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in an Bargaining Model." *Econometrica* 50:97–110.

Ruckriegel, Karlheinz. 2010. “Die Wiederentdeckung des Menschen in der Ökonomie.” In *Vortrag im Rahmen der Tagung "Ökonomische Rationalitätenkaskade" an der Universität Koblenz-Landau, überarbeitete Fassung vom 30.12.*

Rudolf, Stefan. 2011. *Militaerorganisatorische und gesellschaftliche Herausforderungen bei der "Transformation" der Bundeswehr in eine Einsatzarmee*. GRIN Verlag.

Ruffle, Bradley J. and Richard H. Sosis. 2006. “Cooperation and the In-Group-Out-Group Bias: A Field Test on Israeli Kibbutz Members and City Residents.” *Journal of economic behavior and organization* 60:147–163.

Russ, Steve. 2011. “The unreasonable effectiveness of mathematics in the natural sciences.” *Interdisciplinary Science Reviews* 36:209–213.

Rutstroem, E. Elisabet and Melonie B. Williams. 2000. “Entitlements and fairness: an experimental study of distributive preferences.” *Journal of Economic Behavior and Organization* 43:75–89.

Sachs, Jeffrey D. 2012. “Government, Geography, and Growth - The True Drivers of Economic Development.” *Foreign Affairs* 91.

Saijo, T. and H. Nakamura. 1995. “The Spite Dilemma in Voluntary Contribution Mechanism Experiments.” *Journal of Conflict Resolution* 39:535–560.

Salant, Stephen W. and Eban Goodstein. 1990. “Predicting Committee Behavior in Majority Rule Voting Experiments.” *The RAND Journal of Economics* 21:293–313.

Samuelson, W. and W. Zeckhauser. 1988. “Status Quo Bias in Decision Making.” *Journal of Risk and Uncertainty* 1:7–59.

Sauermann, Jan. 2012. “Ticken Gruppen anders? Ein Laborexperiment zur unterschiedlichen Motivation von Individuen und Gruppen in Verhandlungen.” In *Jahrbuch für Handlungs- und Entscheidungstheorie*, edited by Thomas Bräuninger, Bächtiger, and Susumu Shikano, number Band 7: Experiment und Simulation, chapter 5, pp. 81–104. VS Verlag für Sozialwissenschaften, Wiesbaden.

Sauermann, Jan and Ulrich Glasmann. 2011. “Entscheidungskosten und Gemeinwohleffekte demokratischer Abstimmungsregeln - eine experimentelle Untersuchung.” *Politische Vierteljahresschrift (PVS)* 52:373–398.

Sauermann, J. and A. Kaiser. 2010. “Taking Others into Account: Self-Interest and Fairness in Majority Decision Making.” *American Journal of Political Science* 54:667–685.

Saunders, Ben. 2010a. “Democracy, Political Equality, and Majority Rule.” *Ethics, The University of Chicago Press* 121:148–177.

Saunders, Ben. 2010b. “Why Majority Rule Cannot Be Based Only on Procedural Equality.” *Ratio Juris* 23:113–122.

Scharpf, Fritz W. 2000. *Interaktionsformen. Akteurzentrierter Institutionalismus in der Politikforschung*. Opladen: Leske and Budrich.

Schäfers, Michael and Joachim Zimmermann. 2005. *Partizipation: Von Solidarität, Eigenverantwortung und bürgerschaftlichem Engagement [Taschenbuch]*. LIT Verlag Münster.

Schmidt, Klaus M. 2010. "Social Preferences and Competition." *Discussion Papers in Economics, University of Munich* 6.

Schmidt, M. G. 2012. *Der deutsche Sozialstaat. Geschichte und Gegenwart*. München: C.H.Beck.

Schnapp, Kai-Uwe, Nathalie Behnke, and Joachim Behnke. 2009. *Datenwelten - Datenerhebung und Datenbestaende in der Politikwissenschaft*. Nomos.

Schneider, G., D. Finke, and S. Bailer. 2010. "Bargaining Power in the European Union: An Evaluation of Competing Game-Theoretic Models." *Political Studies* 58:85–103.

Schneider, Gerald, Bernard Steunenberg, and Mika Widgren. 2006. "Evidence with insigt: What models contribute to EU research." In *The European Union Decides. The Empirical Relevance of Policy Making Models*, edited by R. Thomson, F. N. Stokman, C. H. Achen, and T. König, chapter 11, pp. 299–316. Cambridge University Press.

Schoenberg, Ronald. 2000. "Constrained Optimization." *GAUSS documentation* http://econ.la.psu.edu/computing/pdf/co.pdf (accessed March 2, 2011).

Schofield, Norman. 1996. *Collective Decision-Making: Social Choice and Political Economy*. Springer, 1 edition.

Schofield, Norman. 2008. *The spatial model of politics*. Routledge, London and New York.

Schopler, John and Chester A. Insko. 1992. "The Discontinuity Effect in Interpersonal and Intergroup Relations: Generality and Mediation." In *European Review of Social Psychology*, edited by Wolfgang Stroebe and Miles Hewstone, pp. 121–151. Chichester: Wiley.

Schotter, Andrew. 2006. "Strong and Wrong: The Use of Ratio-nal Choice Theory in Experimental Economics." *Journal of Theoretical Politics* 18:498–511.

Schram, A. 2005. "Artificiality: The tension between internal and external validity in economic experiments." *Journal of Economic Methodology* 12:225–237.

Schram, Arthur J.H.C. 2002. "Experimental Public Choice." *Tinbergen Institute Discussion Papers* 106.

Schuman, Howard and Stanley Presser. 1996. *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Sage Publications.

Schwartz, Thomas. 1977. "Collective Choice, Separation of Issues and Vote Trading." *The American Political Science Review* 71:999–1010.

Selten, Reinhard and Rosemarie Nagel. 1998. "Das Zahlenwahlspiel-Hintergruende und Ergebnisse." *Spektrum der Wissenschaft* February:16–22.

Sen, A. 1995. "Rationality and Social Choice." *American Economic Review* 85:1–24.

Sen, Amartya. 1997. "Maximization and the Act of Choice." *Econometrica* 65 (4):745–780.

Sen, A. K. 1977. "Rational Fools a critique of the behavioural foundations of economic theory." *Philosophy and Public Affairs* 6:317–344.

Sengul, Metin, Javier Gimeno, and Jay Dial. 2012. "Strategic Delegation: A Review, Theoretical Integration, and Research Agenda." *Journal of Management* 38:375–414.

Severinov, Sergei. 2008. "The value of information and optimal organization." *RAND Journal of Economics* 39:238–265.

Shadish, William R., Thomas D. Cook, and Donald Thomas Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.

Shapiro, S. S. and M. B. Wilk. 1965. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika* 52:591–611.

Shavit, T., D. Sonsino, and U. Benzion. 2001. "A comparative study of lotteries-evaluation in class and on the web." *Journal of Economic Psychology* 22:483–491.

Shepard, Roger N. 1991. "Integrality versus separability of stimulus dimension: From an early convergence of evidence to a proposed theoretical basis." In *The perception of structure: Essays in honor of Wendell R. Garner.*, edited by Gregory R. Lockhead and James R. Pomerantz, pp. 53–71. Washington, DC: American Psychological Association.

Sheppard, Simon G. 1998. *Game-Theoretic Models of Human Behaviour*. Heretical Press.

Shepsle, Kenneth A. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science* 23:27–59.

Shepsle, Kenneth A. and Mark S. Bonchek. 1997. *Analyzing politics : rationality, behavior, and institutions*. New York: W.W. Norton.

Shepsle, Kenneth A. and Gary W. Cox. 2007. "Majority Cycling and Agenda Manipulation: Richard McKelvey's Contributions and Legacy." In *Positive Change in Political Science: The Legacy of Richard McKelvey's Most Influential Writings*, edited by John Aldrich, James E. Alt, and Arthur Lupia. Ann Arbor: University of Michigan Press.

Shepsle, Kenneth A. and Barry R. Weingast. 1984. "Uncovered Sets and Sophisticated Voting Outcomes with Implications for Agenda Institutions." *American Journal of Political Science* 28:49–74.

Shepsle, Kenneth A. and Barry R. Weingast. 1995. *Positive theories of congressional institutions*. Ann Arbor: University of Michigan Press.

Shepsle, Kenneth A. and Barry R. Weingast. 2004. "Structure-induced equilibrium and legislative choice." *Public Choice* 37 (3):503–519.

Sherstyuk, Katerina, Tarui Nori, and Saijo Tatsuyoshi. 2011. "Payment schemes in infinite-horizon experimental games." *Working Papers Series, University of Hawai'i* .

Shikano, Susumu. 2009. "Simulating party competition and vote decision under mixed member electoral systems." *AUCO Czech Economic Review* 3:270–291.

Shils, Edward A. and Henry A. Finch. 1949. *Max Weber on the Methodology of the Social Sciences*. New York: Free Press.

Shu, Min. 2003. "Nonseparable Preferences, Public Opinion and Referendums on European Integration." *mimeo. Paper presented at the ECPR General Conference, Marburg, Germany.* http://www.aoni.waseda.jp/min.shu/paper.html (accessed December 8, 2012).

Shuttleworth, Martyn. 2008. "Double Blind Experiment." *Experiment Resources* http://www.experiment-resources.com/double-blind-experiment.html (accessed January 13, 2012).

Shy, Oz. 1996. *Industrial Organization: Theory and Applications*, volume 1. The MIT Press, 1st edition.

Siegel, Sidney. 1988. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Siegel, Sidney. 2001. *Nichtparametrische statistische Methoden*. Eschborn bei Frankfurt a.M.: Klotz, 5., unveränd. aufl. edition.

Signorino, Curtis S. 1999. "Strategic interaction and the statistical analysis of international conflict." *American Political Science Review* 93:279–297.

Signorino, Curtis S. 2003. "Structure and Uncertainty in Discrete Choice Models." *Political Analysis* 11:316–344.

Signorino, Curtis S. and Kuzey Yilmaz. 2003. "Strategic Misspecification in Regression Models (with . . 47(3):." *American Journal of Political Science* 47:551–566.

Siminski, Peter. 2008. "Order Effects in Batteries of Questions." *Quality and Quantity: International Journal of Methodology* 42:477–490.

Simon, D., D. C. Krawczyk, and K. J. Holyoak. 2004. "Construction of preferences by constraint satisfaction." *Psychological Science* 15:331–336.

Simon, Herbert. 1957. "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics* 59:99–118.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106.

Simon, Herbert A. 1982. *Models of bounded rationality*. Cambridge, Mass. ; London: MIT Press.

Simon, Herbert A. 1993. *Homo rationalis. Die Vernunft im menschlichen Leben.* Frankfurt a.M.: Campus.

Simon, Herbert A., Massimo Egidi, and Robin Lapthorn Marris. 1992. *Economics, bounded rationality and the cognitive revolution*. Aldershot: Elgar.

Slapin, Jonathan and Sven Oliver Procksch. 2010. "Look who's talking: Parliamentary debate in the European Union." *European Union Politics September* 11:333–357. erste Quelle war 2008, nicht 2010.

Slapin, J. B. and S. O. Proksch. 2008. "A scaling model for estimating time-series party positions from texts." *American Journal of Political Science* 52:705–722.

Smirnov, O. 2009. "Endogenous choice of amendment agendas: types of voters and experimental evidence." *Public Choice* 141:277–290.

Smith, A. 1759. *The Theory of Moral Sentiment*. London: A. Millar. P. 1: How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London.

Smith, Charles E., Robert D. Brown, John M. Bruce, and L. Marvin Overby. 1999. "Party Balancing and Voting for Congress in the 1996 National Election." *American Journal of Political Science* 43:737–64.

Smith, Vernon L. 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy, University of Chicago Press* 70:111–137.

Smith, Vernon L. 1964. "Effect of Market Organization on Competitive Equilibrium." *Quarterly Journal of Economics* 78:182–205.

Smith, Vernon L. 1973. "Notes on Some Literature in Experimental Economics." *Social Science working paper California Institut of Technology* 21.

Smith, Vernon L. 1976. "Experimental economics: Induced value theory." *The American Economic Review* 65(2):274–279.

Smith, Vernon L. 1991. *Papers in Experimental Economics*. Cambridge University Press.

Smith, Vernon L. and James M. Walker. 1993. "Money Rewards and Decision Costs in Experimental Economics." *Economic Inquiry* 15:245–261.

Smithies, Arthur. 1941. "Optimum Location in Spatial Competition." *Journal of Political Economy, University of Chicago Press* 49:423–439.

Snedecor, G. W. and W. G. Cochran. 1989. *Statistical Methods*. Ames, IA: Iowa State University Press, 8th edition.

Sobel, J. 2005. "Interdependent preferences and reciprocity." *Journal of Economic Literature* 43:392–436.

Soberg, M. S. and M. P. Wattenberg. 2001. *Mixed-member electoral systems: the best of both worlds?* Oxford, Oxford University Press.

Somers, Robert H. 1962. "A New Asymmetric Measure of Association for Ordinal Variables." *American Sociological Review* 27:799–811.

Soto, Fabian A. and Edward A. Wasserman. 2010. "Integrality and Separability of Stimulus Dimensions and Multidimensional Generalization in Pigeons." *Journal of Experimental Psychology: Animal Behavior Processes* 36:194–205.

Sprinthall, Richard C. 2011. *Basic Statistical Analysis*. Prentice Hall, 9 edition.

Sribney, Willim. 1999. "What does "completely determined" mean in my logistic regression output?" *Statalist, Stata 10 FAQ, StataCorp LP.* http://www.stata.com/support/faqs/statistics/completely-determined-in-logistic-regression (accessed April 4, 2013).

Stahl, D. O. and E. Haruvy. 2008. "Subgame perfection in ultimatum bargaining trees." *Games and Economic Behavior* 63:292–307.

Stanovich, K. E. and R. F. West. 2000. "Individual differences in reasoning: Implications for the rationality debate?" *Behavioral and Brain Sciences* 23:645.

Steunenberg, B. 1994. "Decision-Making under Different Institutional Arrangements - Legislation by the European-Community." *Journal of Institutional and Theoretical Economics-Zeitschrift Fur Die Gesamte Staatswissenschaft* 150:642–669.

Steunenberg, Bernard and Torsten J. Selck. 2006. "Testing Procedural Models of European Union Legislative Decision-Making." In *The European Union Decides*, edited by Christopher H. Achen Robert Thomson, Frans N. Stokman and Thomas König, pp. 54–84. Cambridge University Press.

Steunenberg, Bernard and Frans van Vught. 1997. *Political institutions and public policy : perspectives on European decision making*. Dordrecht ; Boston: Kluwer Academic Publishers.

Stigler, G. G. 1950. "The development of utility theory, I,II." *Journal of Political Economy* I, II:43, 307–327 373–396.

Stiglitz, Joseph E. 1972. "Some Aspects of the Pure Theory of Corporate Finance: Bankruptcies and Take-Overs." *Bell Journal of Economics, The RAND Corporation* 3:458–482.

Stokman, Frans N. and Bruce Bueno De Mesquita. 1994. *European Community Decision Making: Models, Applications, And Comparisons*. New Haven: Yale University Press.

Stokman, Frans N. and Reinier Van Oosten. 1994. "The Exchange of Voting Positions: An Object-Oriented Model of Policy Networks." In *European Community Decision Making: Models, Applications and Comparisons*, edited by Bruce Bueno de Mesquita and Frans N. Stokman, pp. 105–127. New Haven: Yale University Press.

Stoll, Heather. 2013. *Changing Societies, Changing Party Systems*. Cambridge University Press.

Stoll, Heather M. 2004. *Social cleavages, political institutions and party systems: putting preferences back into the fundamental equation of politics*. Dissertation, Stanford University, Deptartment of Political Science.

Strom, Gerald Steven. 1990. *The logic of lawmaking: a spatial theory approach*. Baltimore: Johns Hopkins University Press.

Strom, Kaare, Wolfgang C. Mueller, and Torbjoern Bergman. 2003. *Delegation and Accountability in Parliamentary Democracies*. Oxford , New York.

Studenmund, A. H. 2006. *Using Econometrics: A Practical Guide*. Pearson: Addison-Wesley.

Sugden, Robert. 1989. "Spontaneous Order." *Journal of Economic Perspectives* III:85–98.

Sunitiyoso, Yos, Erel Avineri, and Kiron Chatterjee. 2011. "The effect of social interactions on travel behaviour: An exploratory study using a laboratory experiment." *Transportation Research Part A: Policy and Practice* 45:332–344.

Suri, S. and D. J. Watts. 2011. "Cooperation and Contagion in Web-Based, Networked Public Goods Experiments." *PLOS one* 6.

Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday.

Sutter, Matthais. 2002. "Public Bad Prevention by Majority Voting on Redistribution - Experimental Evidence." *Group Decision and Negotiation* 11:415–428.

Sutter, Matthias. 2008. "Individual behavior and group membership: Comment." *Jena Economic Research Papers, Max-Planck-Institute of Economics* 075.

Swedberg, Richard and Andreas Maurer. 2009. *Grundlagen der Wirtschaftssoziologie*. Verlag für Sozialwissenschaften.

Taleb, Nassim Nicholas. 2007. *The black swan - The impact of the highly improbable*. Random House.

Taylor, Michael. 2006. *Rationality and the Ideology of Disconnection*. New York: Cambridge University Press.

Tegmark, Max. 2007. "The Mathematical Universe." *Foundations of Physics* 38:101–150. Dept. of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139.

Thaler, Richard H. 1980. "Toward a positive theory of consumer choice." *Journal of Economic Behavior and Organization* 1:39–60.

Thaler, Richard H. and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness*. Penguin.

Thiem, Janina. 2009. *Nationale Parteien im Europaeischen Parlament: Delegation, Kontrolle und politischer Einflus*. VS Verlag für Sozialwissenschaften, Wiesbaden.

Thomson, Robert. 2011. *Resolving controversy in the European Union*. Cambridge University Press.

Thomson, R., J. Arregui, D. Leuffen, R. Costello, J. Cross, R. Hertz, and T. Jensen. 2012. "A new dataset on decision-making in the European Union before and after the 2004 and 2007 enlargements (DEUII)." *Journal of European Public Policy* 19(4):604–622.

Thomson, R., F. N. Stokman, C. H. Achen, and T. König. 2002. "Report on Commission proposal COM(1999)582." *mimeo. DEU field report.*

Thomson, R., F. N. Stokman, C. H. Achen, and T. König. 2006. *The European Union decides*. Cambridge, UK ; New York: Cambridge University Press.

Thye, Shane R. 2007. "Logical and Philosophical Foundations of Experiments in the Social Sciences." In *Laboratory Experiments in the Social Sciences*, edited by Murray Webster and Jane Sell, chapter 3, pp. 57–86. Elsevier: London.

Tirole, Jean. 1988. *The theory of industrial organization*. Cambridge, Mass.: MIT Press.

Tolwinski, B. 2003. *A concept of cooperative equilibrium for dynamic games*. Warsaw, Poland: Polish Academy of Sciences.

Torgerson, Warren S. 1952. "Multidimensional scaling: Theory and Method." *Psychometrica* 17:401–419.

Torgerson, Warren S. 1967. *Theory and methods of scaling*. New York [u.a.]: Wiley.

Tornay, S.C. 1938. *Ockham: Studies and Selections*. La Salle, IL: Open Court.

Traub, Stefan, Christian Seidl, and Ulrich Schmidt. 2009. "An experimental study on individual choice, social welfare, and social preferences." *European Economic Review* 53:385–400.

Tremewan, James, Chloe Le Coq, and Alexander K. Wagner. 2012. "Social Preferences and Bounded Rationality in the Centipede Game." *mimeo. Department of Economics, University of Vienna.* http://vimeo.com/37037476 (accessed December 12, 2012).

Tsebelis, G. 1994. "The Power of the European Parliament as a Conditional Agenda Setter." *American Political Science Review* 88:128–142.

Tsebelis, George. 2002. *Veto players : how political institutions work*. Princeton, N.J.: Princeton University Press.

Tsebelis, G. and G. Garrett. 2001. "The institutional foundations of intergovernmentalism and supranationalism in the European Union." *International Organization* 55:357.

Tullock, Gordon. 1981. "Why so much stability." *Public Choice* 37:189–204.

Tversky, Amos and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211:453–458.

Tversky, Amos and Daniel Kahneman. 1986. "Rational Choice and the Framing of Decisions." *The Journal of Business* 59:251–278.

Tversky, Amos and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *The Quarterly Journal of Economics* 106:1039–1061.

van de Kragt, Alphons J. C., John M. Orbell, and Robyn M. Dawes. 1983. "The Minimal Contributing Set as a Solution to Public Goods Problems." *The American Political Science Review* 77(1).

van Deemen, Adrian. 2009. "Structure-induced (Dis)Equilibrium." *mimeo. Institute for Management Research, University of Nijmegen.* http://www.gate.cnrs.fr/IMG/pdf/van_Deemen.pdf (accessed October 29, 2011).

Van den Bos, Jan M. M. 1991. *Dutch EC policy making: a model-guided approach to coordination and negotiation*. Amsterdam: Thesis Publishers.

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil. 1991. "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure." In *Average Opinion Games*, volume 106, pp. 885–910. Quarterly Journal of Economics.

Vanberg, Christoph. 2008. "Why Do People Keep Their Promises? An Experimental Test of Two Explanations." *Econometrica* 76:1467–1480.

Vanberg, Viktor. 2002. "Rational Choice vs. Program-Based Behavior: Alternative Theoretical Approaches and Their Relevance for the Study of Institutions." *Rationality and Society* 14:7–54.

Veblen, Thorstein. 1899. *The theory of the leisure class; an economic study in the evolution of institutions*. London: Macmillan & co., ltd.

Velten, Carlo and Steve Janata. 2012. *Datenexplosion in der Unternehmens-IT: Wie Big Data das Business und die IT verändert*. Eine Studie der Experton Group AG im Auftrag der BT (Germany) GmbH & Co. oHG.

Velupillai, K. Vela. 2005. "The unreasonable ineffectiveness of mathematics in economics." *Cambridge Journal of Economics* 29:849–872.

Vincent, Leah-Martin. 2012. "The Role of Gender and Race in Cooperation and Trust: Experimental Examination of Identity Based Strategic Interaction." *mimeo. University of California, San Diego.* http://www.econ.ucsd.edu/~rstarr/191AFall2012/191%20Final%20Paper%20-%20Leah-Martin%20%282012%29.pdf (accessed February 2, 2013).

von Bismarck, Otto. 1863. "Speech to the Herrenhaus at December 18." In *Oxford Dictionary of Quotations*. 3d, p. 84 edition.

von der Leyen, Ursula. 2012. "Rede von Ursula von der Leyen, Bundesministerin fuer Arbeit und Soziales, anlaesslich der Senkung des Rentenbeitragssatzes 2013 im Deutschen Bundestag." *Haushaltsrede* 27.09. 13:50.

von Neumann, John and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, NJ. Princeton University Press, 1st edition. 1944, sec.ed. 1947, th.ed. 1953.

von Prittwitz, Volker. 2007. *Vergleichende Politikanalyse*. Utb, Stuttgart.

von Weizsäcker, Carl Christian. 2012. "Eine Makro- Perspektive der Evolutionsökonomie: Das Problem der Vorsorge." *Vortrag Evolutionsökonomischer Ausschuss Kloster Beuron 13. Juli 2012* http://www.denkwerkzukunft.de/downloads/CCvW-Beuron-Vortrag-Jul%202012.pdf (accessed August 22, 2012).

Vredenburgh, Donald J. and John G. Maurer. 1984. "A Process Framework of Organizational Politics." *Human relations* 37:47–65.

Wachendorfer-Schmidt, Ute. 2005. *Politikverflechtung im vereinigten Deutschland*. VS Verlag für Sozialwissenschaften, Wiesbaden.

Wagschal, Uwe. 2009. *Statistik fuer Politikwissenschaftler*. Muenchen: Oldenbourg, R.

Wallenius, J., J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, and K. Deb. 2008. "Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead." *Management Science* 54:1336–1349.

Walliser, Bernard. 2008. *Cognitive economics*. Berlin: Springer.

Walter, Stephen D., Alvan R. Feinstein, and Carolyn K. Wells. 1987. "Coding Ordinal Independent Variables in Multiple Regression Analyses." *American Journal of Epidemiology* 125:319–23.

Wandling, Richard. 2012. "Rationality and Rational Choice." In *21st Century Political Science*, edited by John Ishiyama and Marijke Breuning, chapter 5, pp. 34–42. Thousand Oaks: Sage.

Wantchekon, L. 2003. "Clientelism and voting behavior - Evidence from a field experiment in Benin." *World Politics* 55:399–422.

Weber, E. U. and E. J. Johnson. 2008. "Decisions under uncertainty: Psychological, economic, and neuroeconomic explanations of risk preference." In *Neuroeconomics: Decision making and the brain*, edited by P. Glimcher, C. Camerer, E. Fehr, and R. Poldrack, chapter 10, pp. 127–144. New York: Elsevier.

Weber, Max. 1968. *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen.

Webster, Murray and Jane Sell. 2007. *Laboratory experiments in the social sciences*. Amsterdam ; Boston: Academic Press/Elsevier: London.

Weibull, Joergen W. 1995. *Evolutionary game theory*. Cambridge, Mass ; London: MIT Press.

Weihe, Anne Cordelia, Tanja Pritzlaff, Frank Nullmeier, Tilo Felgenhauer, and Britta Baumgarten. 2008. "Wie wird in politischen Gremien entschieden? Konzeptionelle und methodische Grundlagen der Gremienanalyse." *Politische Vierteljahresschrift (PVS)* 49:339–359.

Weingast, Barry R. 2009. "Second generation fiscal federalism: The implications of fiscal incentives." *Journal of Urban Economics* 64:279–293.

Weise, Frank-Juergen, Hans Heinrich Driftman, Hans-Ulrich Klose, Juergen Kluge, Karl-Hein Lather, and Hedda von Wedel. 2010. "Vom Einsatz her Denken - Konzentration, Flexibilitaet, Effizienz." Technical report, Bericht der Strukturkomission der Bundeswehr.

Wicksell, Knut. 1896. *Finanztheoretische Untersuchungen: nebst Darstellung und Kritik des Steuerwesens Schwedens*. Jena: Fischer Verlag.

Wigner, Eugene. 1960. *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, volume 13 of *Communications in Pure and Applied Mathematics*. New York: John Wiley & Sons.

Wilcoxon, Frank. 1945. "Individual comparisons by ranking methods." *Biometrics Bulletin* 1:80–83.

Wildschut, Tim, Hein F. M. Lodewijkx, and Chester A. Insko. 2001. "Toward a Reconciliation of Diverging Perspectives on Interindividual-Intergroup Discontinuity: The Role of Procedural Interdependence." *Journal of Experimental Social Psychology* 37:273–285.

Wildschut, Tim, Brad Pinter, Jack L. Vevea, Chester A. Insko, and John Schopler. 2003. "Beyond the Group Mind: A Quantitative Review of the Interindividual-Intergroup Discontinuity Effect." *Psychological Bulletin* 129:698–722.

Williamson, Oliver E. 1975. *Markets and hierarchies, analysis and antitrust implications : a study in the economics of internal organization*. New York: Free Press.

Willits, Fern K. and Bin Ke. 1995. "Part-Whole Question Order Effects: Views of Rurality." *The Public Opinion Quarterly* 59:392–403.

Wilson, E. B. 1927. "Probable inference, the law of succes-sion, andstatistical inference." *Journal of the American Statistical Association* 22:209–212.

Wilson, Rick and Catherine Eckel. 2011. "Trust and Social Exchange." In *The Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, chapter 17. New York: Cambridge Unviersity Press.

Wilson, R. K. 2007. "Voting and Agenda Setting in Political Science and Economics." In *Laboratory Experiments in the Social Sciences*, edited by M. Webster and J. Sell, pp. 433–458. San Diego, CA: Elsevier.

Witt, L. A., Martha C. Andrews, and K. Michele Kacmar. 2000. "The Role of Participation in Decision-Making in the Organizational Politics-Job Satisfaction Relationship." *Human relations* 53:341–358.

Wolf, Frieder. 2010. "Enlightened Eclecticism or Hazardous Hodgepodge? Mixed Methods and Triangulation Strategies in Comparative Public Policy Research." *Journal of Mixed Methods Research* 4:144–167.

Würzner, Eckart. 2010. "Das Signal ist angekommen - Interview mit Oberbürgermeister Dr. Eckart Würzner zur Stadthallen-Erweiterung." *Rhein-Neckar-Zeitung* 15.06.

Xia, Lirong, Vincent Conitzera, and Jerome Lang. 2011. "Strategic Sequential Voting in Multi-Issue Domains and Multiple-Election Paradoxes." *Preprint submitted to Journal of Economic Theory* .

Yezer, Tony, Robert Goldfarb, and Paul Poppen. 1996. "Does Studying Economics Discourage Cooperation?: Watch What We Do, Not What We Say or How We Play." *Journal of Economic Perspectives* 10:171–186.

Zaller, J. and S. Feldman. 1992. "A Simple Theory of the Survey Response - Answering Questions Versus Revealing Preferences." *American Journal of Political Science* 36:579–616.

Zizzo, Daniel. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics* 13:75–98.

Zizzo, Daniel John and Jonathan H. W. Tan. 2009. "Game Harmony: A Behavioral Approach to Predicting Cooperation in Games." *ICBBR Working Papers, International Centre for Behavioural Business Research* 4.

Zohlnhöfer, Reimut. 2008. "Foederalismusreform und die Entwicklung der Zustimmungsbeduerftigkeit von Bundesgesetzen: Versuch einer Klaerung." *Zeitschrift fuer Parlamentsfragen* 39:415–419.

Zohlnhöfer, Reimut. 2009. "Der Politikverflechtungsfalle entwischt? Die Effekte der Foederalismusreform I auf die Gesetzgebung." *Zeitschrift fuer Politikwissenschaft* 19. Jahrgang:39–76.

Zuckerman, Marvin. 1994. *Behavioural expressions and biosocial bases of Sensation Seeking*. Cambridge: Cambridge University Press.

Appendix

A.1 Software used

GAUSS

In Sec. 6.4 I use GAUSS' (Version 9.0.2, build 1114) constrained optimization algorithm[453] (Schoenberg, 2000) to estimate a random-utility mixture model for an individual's choice determinants. This follows the approach of Goeree et al. (2002) but adds sincere and sophisticated voting types. In situations where all subjects vote sincerely this task could be accomplished by using conditional logistic regression which implies a "logit equilibrium" (Goeree et al., 2002, p. 262). When including sophisticated voting the corresponding choice probabilities are adjusted according to the equilibrium solution concept.

STATA

STATA (Version 12.0) serves as main computational tool for the multiple calculations in this study. Most of the commands used can be categorized as belonging to the standard tool box. Not quite ordinary are the application of a local polynomial smoother which used STATA's *kdensity*[454] estimation in Sec. 3.5 and the implementation of a two-way error components model in Sec. 5.1 by using STATA's *xtmelogit*[455] (multilevel mixed-effects) logistic regression. Also, I apply STATA's *somersd* package[456] in Sec. 5.2 to assess trends over single rounds of my experiment.

MATLAB

I use MATLAB's (Version 7.12.0.635) "Global Optimization Toolbox"[457] (Version 3.2 - Release 2011b) and the related extension "Genetic Algorithm and Direct Search Toolbox"[458] in Sec. 3.5 for simulating the effect of different NSP values. Based on the idea of Perttunen et al. (1993) this derivative-free global optimization algorithm enables me to find the global optimum of the objective function. Due to the discontinuity of the function

[453] Source: `http://www.gaussian.com/g_tech/g_ur/k_opt.htm`.
[454] Source: `http://www.stata.com/help.cgi?kdensity`.
[455] Source: `http://www.stata.com/help.cgi?xtmelogit`.
[456] The *somersd* package can be obtained from the Statistical Software Components (SSC) archive, which is accessible under `http://www.repec.org`. SSC is one of the most important download sites for user-written STATA packages.
[457] Source: `http://www.mathworks.de/products/global-optimization`.
[458] Source: `http://www.mathworks.com/help/releases/R13sp2/pdf_doc/gads/gads_tb.pdf`.

this is impossible while using standard optimization methods for constrained nonlinear optimization problems such as Sequential Quadratic Programming.

R / WORDFISH

In Sec. 7.4 I use the word scaling algorithm WORDFISH[459] (Proksch and Slapin, 2009), written in R statistical language (Version 2.14.2., release date 2012-02-19). The algorithm extracts (spatial) positions from text documents using word frequencies and places different documents into a single dimension of discrepancy by means of maximum likelihood estimations. The scaling technique does not need any anchoring documents to perform the analysis. Instead, it relies on a statistical model of word counts, more precisely a Poisson distribution (Poisson, 1837) of word frequencies.

My analysis also takes advantage of the text mining package already included in R (tm package, Feinerer et al., 2008). This package possesses a stemming algorithm and a stop-word dictionary. A well-documented example of how to apply WORDFISH can be found in Slapin and Proksch (2008).

z-Tree

I implemented my experiment using the z-Tree software (Version 3.3.11) developed by Fischbacher (2007). The acronym stands for Zurich Toolbox for Readymade Economic Experiments. The software facilitates to develop and carry out laboratory experiments as many standards are predefined. The program can also be adjusted to fit a wide range of experimental designs. It requires no previous programming knowledge and can be licensed free of charge. Under `http://www.iew.uzh.ch/ztree/index.php` the current version (3.3.12) and much more information (z-Tree Wiki, FAQ, etc.) can be obtained.

[459] Source: `http://www.wordfish.org`.

A.2 Weighted Euclidean distance including nonseparable preferences

The standard weighted Euclidean distance (cf. Hinich and Munger, 1997) is represented by $WED\,(\theta, x) = \sqrt{(\theta - x)^T\, A\,(\theta\, - x)}$, where in a d-dimensional space $\theta = (\theta_1,\theta_2,\theta_3,...,\theta_d)$ describes an actor's unconditional first preference, $x = (x_1,x_2,x_3,...,x_d)$ describes the policy and $A = \begin{bmatrix} a_{11} & ... & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & ... & a_{dd} \end{bmatrix}$ describes each dimension's salience in its main diagonal $(a_{11}, ..., a_{dd})$ and the conditionality between two dimensions in the secondary diagonals $(a_{d1}, ..., a_{1d})$.

Applying the standard equation to a simple two-dimensional example (dimensions i and j) results in

$$WED\,(\theta, x) = \sqrt{\begin{bmatrix} \theta_i - x_i & \theta_j - x_j \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix} \begin{bmatrix} \theta_i - x_i \\ \theta_j - x_j \end{bmatrix}} \tag{A.1}$$

Eqn. A.1 applies to the case of reciprocal nonseparability. In the case of non-reciprocal nonseparability (cf. Sec. 2.4.5) an actor's conditional allocation dimension j is indiscriminant with respect to the direction towards which the outcome deviates from the unconditional ideal point in dimension i. An increase in the absolute distance between their unconditional ideal policy and the outcome causes a decrease in their conditional allocation preferences. Accordingly, I modify the standard equation by using absolute distance $|\theta_i - x_i|$ on the policy dimension:

$$WED\,(\theta, x) = \sqrt{\begin{bmatrix} |\theta_i - x_i| & \theta_j - x_j \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix} \begin{bmatrix} |\theta_i - x_i| \\ \theta_j - x_j \end{bmatrix}}$$

$$= \sqrt{\begin{bmatrix} a_{ii}\,|\theta_i - x_i| + a_{ij}\,(\theta_j - x_j) & a_{ji}\,|\theta_i - x_i| + a_{jj}\,(\theta_j - x_j) \end{bmatrix} \begin{bmatrix} |\theta_i - x_i| \\ \theta_j - x_j \end{bmatrix}}$$

$$= \sqrt{a_{ii}\,|\theta_i - x_i|\,|\theta_i - x_i| + a_{ij}\,(\theta_j - x_j)\,|\theta_i - x_i| + a_{ji}\,|\theta_i - x_i|\,(\theta_j - x_j) + a_{jj}\,(\theta_j - x_j)\,(\theta_j - x_j)}$$

As it always holds that $|\theta_i - x_i|\,|\theta_i - x_i| = (\theta_i - x_i)^2$ and that $(\theta_j - x_j)\,|\theta_i - x_i| = |\theta_i - x_i|\,(\theta_j - x_j)$ the equation can be simplified to

$$WED(\theta, x) = \sqrt{a_{ii}\,(|\theta_i - x_i)^2 + a_{ij}\,(\theta_j - x_j)\,|\theta_i - x_i| + a_{ji}\,|\theta_i - x_i|\,(\theta_j - x_j) + a_{jj}\,(\theta_j - x_j)^2}$$

and finally to

$$WED(\theta, x) = \sqrt{a_{ii}\,(|\theta_i - x_i)^2 + (a_{ij} + a_{ji})\,(\theta_j - x_j)\,|\theta_i - x_i| + a_{jj}\,(\theta_j - x_j)^2} \tag{A.2}$$

Eqn. A.2 corresponds to Eqn. 2.11 in Sec. 2.4.5.

A.3 Structure of the DEU field reports and expert interviews

The field reports of the DEU project (Thomson et al., 2006) collected information on 66 contested law-making proposals. The reports include the following data, which has been gathered by expert interviews:

- Descriptive information on each proposal as identifying number, legislative procedure, date of introduction and revision,
- Descriptive information on the data collection as date and duration of the interview, name, nationality and profession of the expert, evaluation of the informant's knowledge on the topic, etc.
- The number of single issues within each proposal,
- Position and salience of national governments, the European Commission and the European Parliament on each issue.

Most important is the issue specification of the proposal. This was obtained through the interviews and accompanied with background information obtained, e.g., from the Legislative Observatory[460]. The interviews enabled the determination of ideal position and salience of actors. Tab. A.2 provides an overview on the content of the expert interviews. Although the reports did not ask for NSP in particular they proved to be a very helpful starting point for the investigation into potential nonseparability (Tab. 3.1).

Table A.2: Structure of the DEU expert interviews

INFORMATION OF INTEREST	SAMPLE QUESTIONS AND TASKS
Issue specification	List the single issues contained in proposal X. Specify each single proposal.
Position specification	Visualize each issue continua as a scale from 0 to 100. State the substantive meaning, in terms of a policy position, of as many points on the scale as possible. What are the preferences of the stakeholders regarding each dimension? Why is this issue of more importance to some stakeholders than to others? What arguments do the stakeholders use for their preferences? Did stakeholders shift in preferences and positions?
Contestation	Purpose of Commission proposal? Why are these issues contested? Did stakeholders make threats or promises? What relationships exist between the issues? To what extent are they independent from each other?
Further comments	Features of the decision-making situation not contained in these issues? Additional information which is important?

[460] The Legislative Observatory constitutes the European Parliament's database for monitoring EU decision-making processes. It can be accessed online under `http://www.europarl.europa.eu/oeil/home/home.do`.

A.4 Three-dimensional contour plots

The three graphs in Fig. A.1 and in Fig. A.2 depict the three-dimensional spatial models for different levels of nonseparability. The graphs correspond to the two-dimensional representations of the illustrative case studies in Sec. 3.3. As their counterparts, the plots depict member states' unconditional ideal positions, the SQ ante, the winset for QMV, the actual policy outcome and the model predictions. The "utility hills" indicate the Nash product within the winset to the SQ. Accordingly, the thick lines demarcate the border of the winset. Please note that due to the undulating contour not all information is visible.

Figure A.1: Three-dimensional contour plots for Council regulation COM1999/163

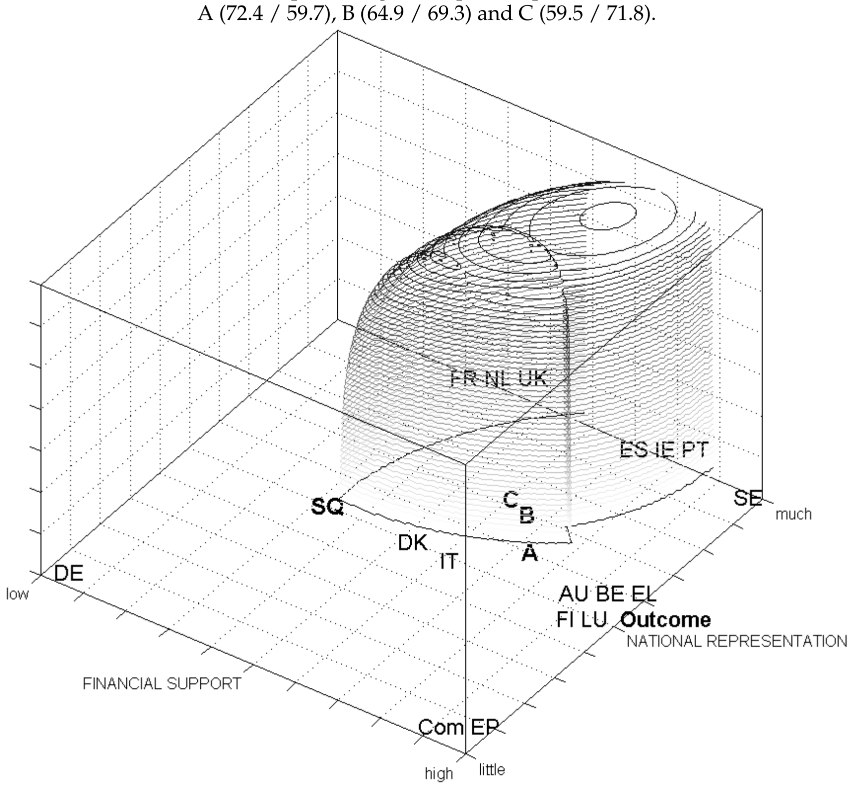
EXPLANATORY NOTE

The three graphs illustrate the Council regulation COM1999/163. The regulation was concerned with the representation of fishery organizations at the EU level. Its two issues were the extent of national representation and the corresponding funding. These two issues form the dimensions of the base area. The utility level is represented as height. Since the utility product does not possess a meaningful unit, I leave the height dimension without index. Instead, the highest peak (or plateau) of the utility hill indicates the maximum relative gain within the Winset. For calculation of the utility gains I assumed a positive but non-reciprocal nonseparability and therefore applied Eqn. 2.11. The SQ was (35/50) and the final outcome (100/50) for this proposal.

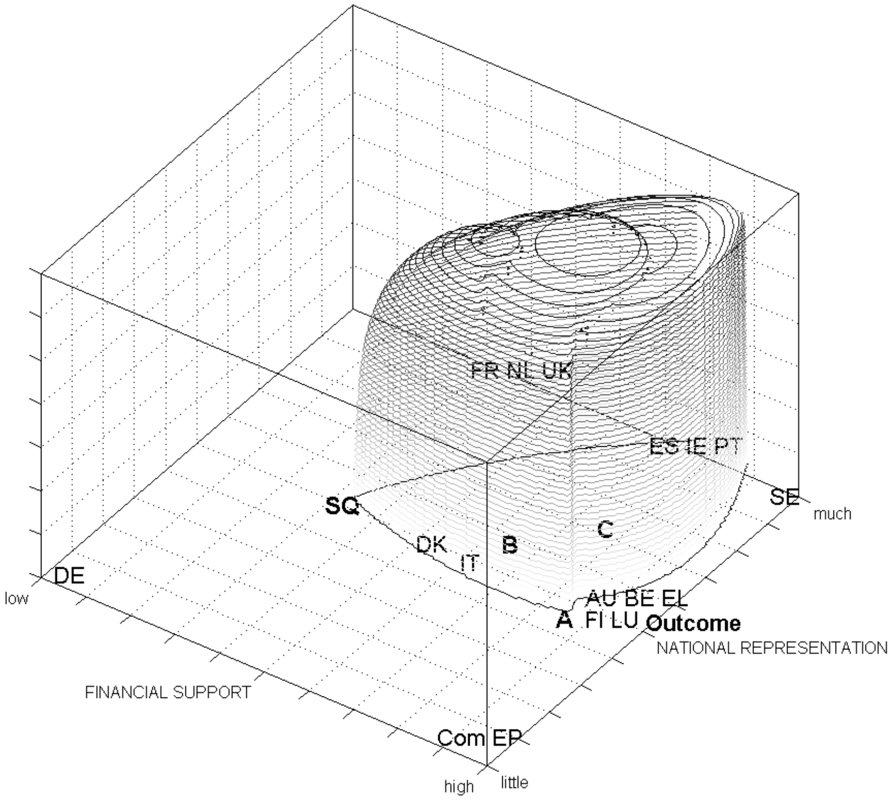
LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom.

Counterpart to Fig. 3.1: separable preferences
A (72.4 / 59.7), B (64.9 / 69.3) and C (59.5 / 71.8).



Counterpart to Fig. 3.2: 50% nonseparable preferences
A (99.9 / 59.7); B (86.7 / 73.1); C (92.0 / 70.5).



Counterpart to Fig. 3.2: 100% nonseparable preferences
A (85.2 / 44.5); B (63.9 / 57.5); C (75.7 / 71.4).

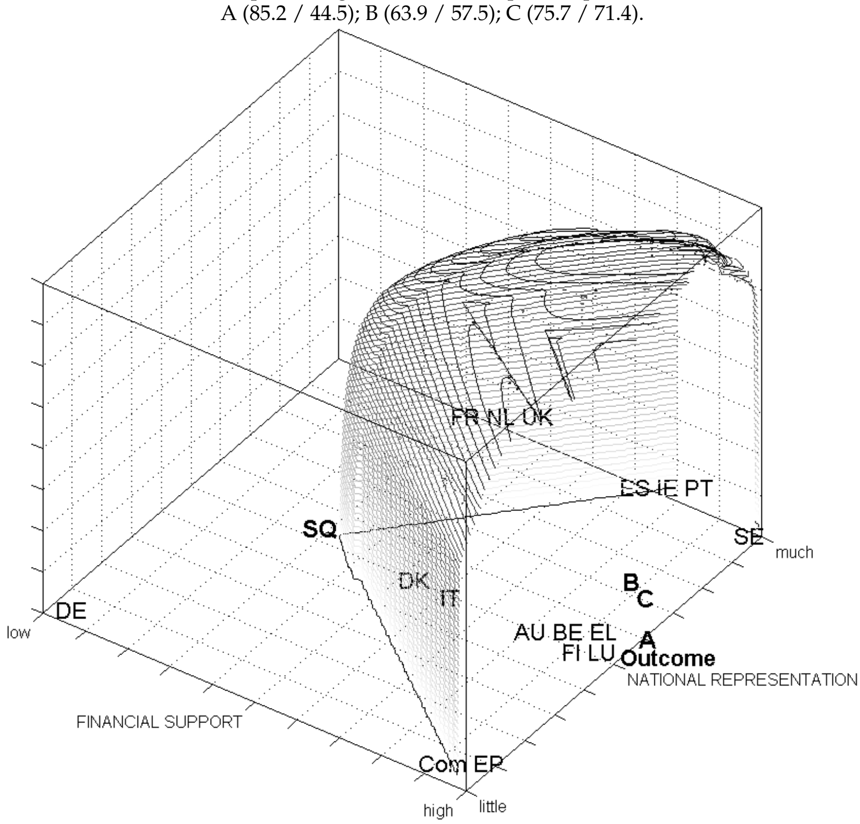


Figure A.2: Three-dimensional contour plots for Council regulation COM1999/582

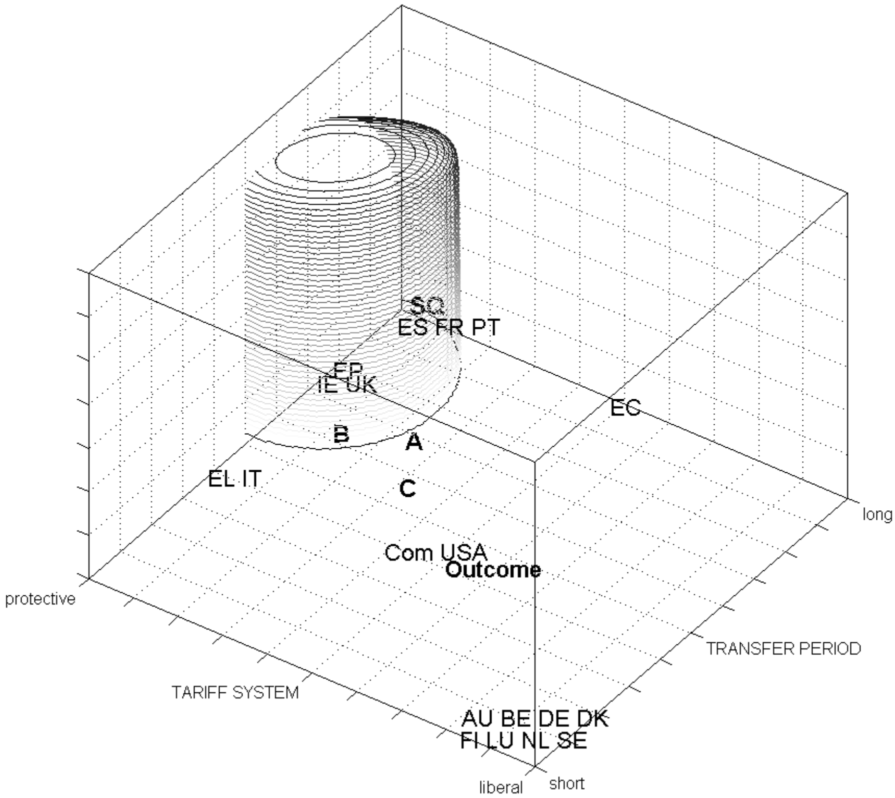
EXPLANATORY NOTE

The three graphs show Council regulation COM1999/582. This regulation amended the organization of the common market in bananas. Its two issues were the type of import regime that would be adopted and the transitional period during which a tariff quota would apply. These two issues form the dimensions of the base area. The utility level is represented as height. Since the utility product does not possess a meaningful unit, I leave the height dimension without index. Instead, the highest peak (or plateau) of the utility hill indicates the maximum relative gain within the Winset. For calculation of the utility gains I assumed mutually positive and reciprocal nonseparability and applied Eqn. 2.9. The SQ was (0/100) and the final outcome (50/40)for this proposal.

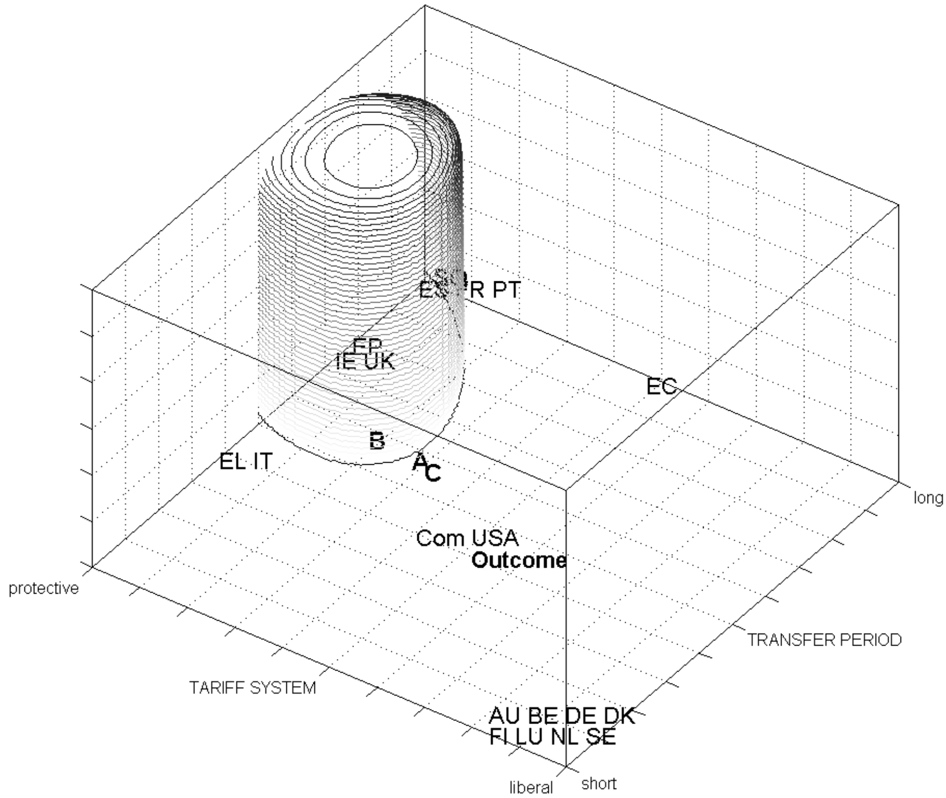
LABELS

A = prediction of the agenda-setting model; B = prediction of the unconstrained bargaining model; C = prediction of the constrained bargaining model; AT = Austria; BE = Belgium; COM = European Commission; DE = Germany; DK = Denmark; EC = Ecuador; EL = Greece; EP = European Parliament; ES = Spain; FI = Finland; FR = France; IE = Ireland; IT = Italy; LU = Luxembourg; NL = Netherlands; PT = Portugal; SE = Sweden; UK = United Kingdom; USA = United States of America.

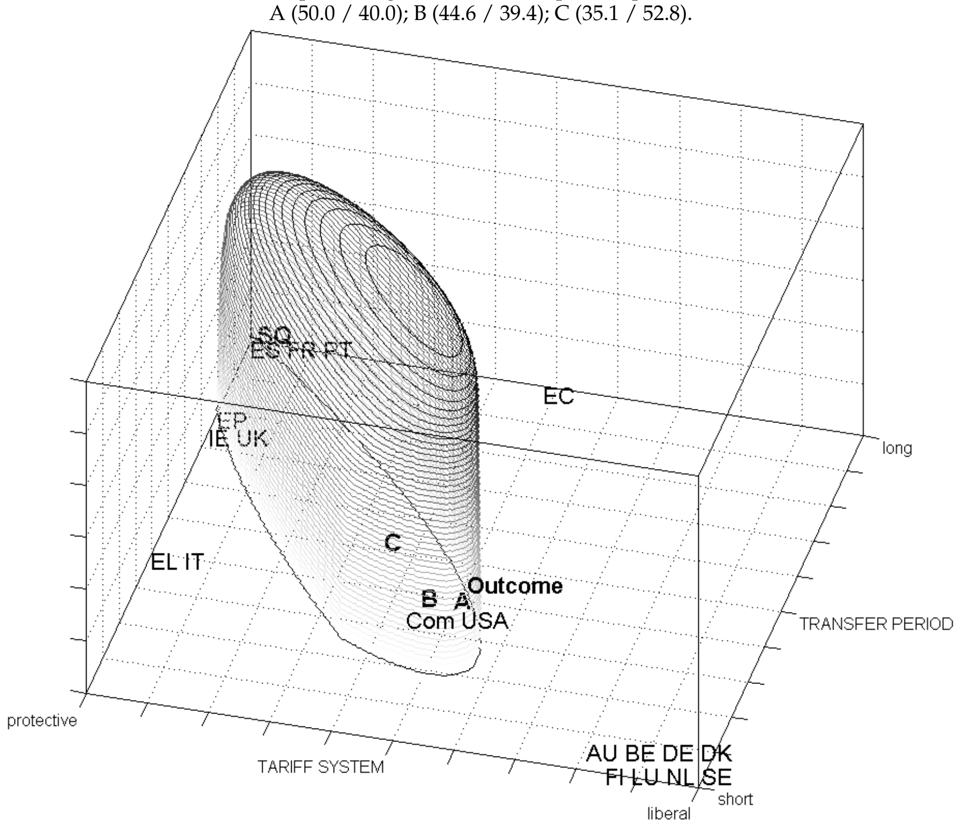
Counterpart to Fig. 3.3: separable preferences
A (24.1 / 68.3), B (11.8 / 62.4) and C (31.3 / 55.8).



Counterpart to Fig. 3.4: 50% nonseparable preferences
A (28.4 / 57.1); B (18.8 / 57.7); C (32.0 / 55.7).



Counterpart to Fig. 3.4: 100% nonseparable preferences
A (50.0 / 40.0); B (44.6 / 39.4); C (35.1 / 52.8).



A.5 The extent of nonseparable preferences in the DEU data set

This section provides the in Chap. 3 discussed NSP coding scheme. It comprises also a short description of every proposal and issue.

Table A.3: The extent of nonseparable preferences in the DEU data set

EXPLANATORY NOTE

The right six columns display the information on nonseparability. In general, it can be read as matrix A, with information on the salience in the main diagonal (s) and information on the presumed direction of nonseparability in the secondary diagonal. A "$+$" indicates a positive and a "$-$" indicates a negative nonseparability between two issues (row and column). The absolute value bars "||" indicate the case of non-reciprocal nonseparability.

Sec. 2.4.5 emphasizes that allocation preferences may be conditionally dependent on the policies pursued but that this must not hold the other way around. In the table, the policy dimension is displayed in the corresponding row, the allocation dimension in the column. For example, when looking at the first proposal in the table below, the preference to prolong or cancel the temporal protection of refugees (issue 1 = "budget allocation") depends on the duration of the temporal protection (issue 2 = "policy 1") and the distribution of asylum seekers across member states (issue 3 = "policy 2").

COM ID	*Short Description*	*Issue Label*	*i*	*i1*	*i2*	*i3*	*i4*	*i5*	*i6*
CNS00127	On minimum standards for giving temporary protection in the event of a mass influx of displaced persons	decision rule to cancel the temporal protection of refugees	1	s					
		duration of the temporal protection	2	+;||	s				
		way in which the asylum seekers have to be distributed among EU member states	3	+;||		s			
CNS00223	Value added tax VAT: length of application of the current minimum standard rate	minimum of VAT standard rate	1	s	+;||				
		time period of measure	2		s				
CNS00250	Proposal for a Council Regulation on the CMO in the Sugar Sector	duration of the extension of the current CMO regime	1	s					
		reduction of production quotas	2	+;||	s	+			
		abolition of subsidies for storage costs	3	+;||	+	s			
CNS00358	Quality strategy for olive oil	extension of the current aid	1	s					
		use and labeling of mixtures of olive oil and vegetable oil	2	+;||	s		+		
		use of talc in the processing of olive oil	3	+;||		s			
		classification of different types of olive oil	4	+;||	+		s		
		labeling of olive oil regarding the country of processing	5	+;||				s	

<table>
<tr><th>COM ID</th><th>Short Description</th><th>Issue Label</th><th>i</th><th>i1</th><th>i2</th><th>i3</th><th>i4</th><th>i5</th><th>i6</th></tr>
<tr><td>CNS98092</td><td>Minimum standards for the protection of laying hens</td><td>use of cages</td><td>1</td><td>s</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>timing for the improvement of the general conditions of the cages</td><td>2</td><td>+;||</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>minimum of cm^2 of cage area for each hen</td><td>3</td><td>+;||</td><td></td><td>s</td><td>+</td><td></td><td></td></tr>
<tr><td></td><td></td><td>timing for the introduction of a compulsory system of cm^2 of cage area for each hen</td><td>4</td><td>+;||</td><td></td><td>+</td><td>s</td><td></td><td></td></tr>
<tr><td></td><td></td><td>timing for banding the cages after the first of January of 2009</td><td>5</td><td>+;||</td><td></td><td></td><td></td><td>s</td><td></td></tr>
<tr><td></td><td></td><td>imported eggs that do not fulfill with the EU rules</td><td>6</td><td>+;||</td><td></td><td></td><td></td><td></td><td>s</td></tr>
<tr><td>CNS98109</td><td>Common organization of the market in beef and veal</td><td>reduction of the intervention price</td><td>1</td><td>s</td><td>-</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>compensation for farmers</td><td>2</td><td>-</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS98110</td><td>Common organization of the market in milk and milk products</td><td>reduction of the intervention price</td><td>1</td><td>s</td><td>-</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>the future of the quota system</td><td>2</td><td>-</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS98299</td><td>EC/Turkey relations: implementation of measures to intensify customs union</td><td>amount of money</td><td>1</td><td>s</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>minority rights</td><td>2</td><td>+;||</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS98347</td><td>Agenda 2000: Financial instrument for fisheries guidance</td><td>size of the scrap-build penalty</td><td>1</td><td>s</td><td>+;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>linkage between the meeting of the objectives of the “Multiannual Guidance Programme” and the allocation of subsidies for fleet renewal and modernization</td><td>2</td><td></td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS98354</td><td>Community action program in the field of civil protection</td><td>duration of the whole program in the field of civil protection</td><td>1</td><td>s</td><td>+;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>amount of money assigned every year for this project</td><td>2</td><td></td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99047</td><td>Common organization of the markets in fishery and aquaculture products</td><td>information to consumers</td><td>1</td><td>s</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>market intervention (subsidies)</td><td>2</td><td></td><td>s</td><td>-</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>trade with third countries (protection)</td><td>3</td><td></td><td>-</td><td>s</td><td></td><td></td><td></td></tr>
<tr><td>CNS99138</td><td>Control measures covered by the Convention on Future Multilateral Co-operation in the North-East Atlantic Fisheries</td><td>implementation of the measures of regional fisheries organizations</td><td>1</td><td>s</td><td>-;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>financing of these measures</td><td>2</td><td></td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99163</td><td>Closer dialogue with the fishing industry and groups affected by the common fisheries policy</td><td>legal basis should be created for expenditure for the support of national level fisheries organizations</td><td>1</td><td>s</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>presence of national representatives in the Advisory Committee on Fisheries</td><td>2</td><td>+;||</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99192</td><td>Setting up an employment committee</td><td>institutional place of the employment committee</td><td>1</td><td>s</td><td>+;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>tasks that should be assigned to the committee</td><td>2</td><td></td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99202</td><td>Council Regulation on production aid for cotton</td><td>level of the Guaranteed National Quantities (GNQ) of Greece and Spain</td><td>1</td><td>s</td><td>-;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>level of the penalties for surplus production</td><td>2</td><td>-;||</td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99214</td><td>Financial and technical measures to accompany</td><td>role of the MED Committee in deciding the projects</td><td>1</td><td>s</td><td>+;||</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>budget assigned by the Commission for financial aid to the MEDA countries</td><td>2</td><td></td><td>s</td><td></td><td></td><td></td><td></td></tr>
<tr><td>CNS99235</td><td>Common organization of the market in bananas</td><td>import system to be adopted for the commerce and distribution of bananas within the EU countries</td><td>1</td><td>s</td><td>+</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>year in which will take into effect the new system for the commerce of bananas within the EU territory</td><td>2</td><td>+</td><td>s</td><td></td><td></td><td></td><td></td></tr>
</table>

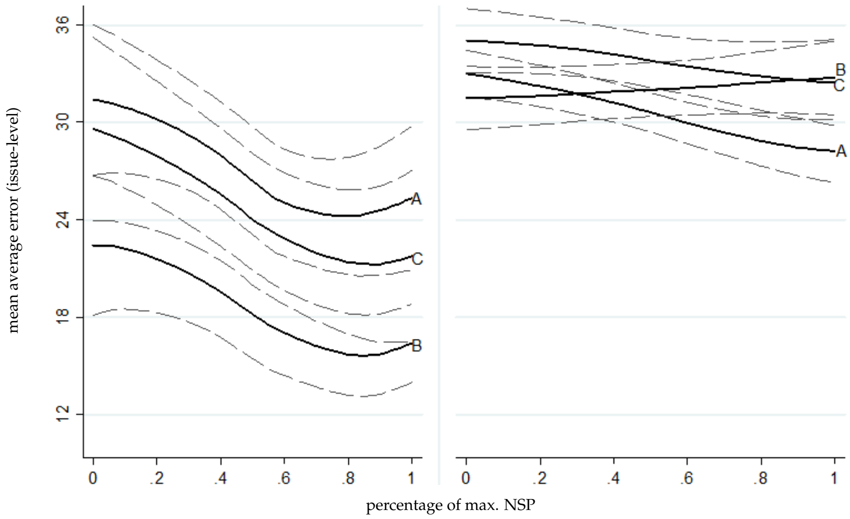
COM ID	Short Description	Issue Label	i	i1	i2	i3	i4	i5	i6
CNS99236	Support system for producers of certain arable crops (flax and hemp)	restriction of support scheme/SQ (long)	1	s	+				
		restriction of support scheme/SQ (short)	2	+	s				
CNS99276	Audio-visual industry: development, distribution and promotion of works	money	1	s					
		distribution	2	+;||	s				
		project	3	+;||		s			
		period	4	+;||			s		
		pilot	5	+;||				s	
COD00032	Public access to documents of the European Parliament, the Council and the Commission	scope of the implementation of the public access to documents	1	s	-	-			
		regimen of exceptions in order to provide public access to documents	2	-	s				
		public access to confidential documents	3	-		s			
		whether the European rules should be applied to National Authorities	4	+;||			s		
COD96085	Intellectual property, original works of art: resale right for the benefit of the author	threshold above which the resale right should apply	1	s		-	+;||		
		cap, or maximum amount of money an artist should receive as their resale right	2		s		+;||		
		degressivity of the resale right	3	-		s	+;||		
		date of implementation	4				s		
COD96112	Directive on coca and chocolate products intended for human consumption	use of vegetable fats other than cocoa butter in chocolate products	1	s			-		
		labeling of chocolate products that contain veg. fats other than cocoa	2	-;||	s				
		derogation for the UK and Ireland regarding milk chocolate	3			s			
		timing of the impact study on developing countries	4	-			s		
COD97359	Directive on the harmonization of certain aspects of copyright and related rights	copyright on internet	1	s					
		time shifting	2	+;||	s				
		exceptions	3	+;||		s			
COD98195	Education, training: Community action program SOCRATES	money	1	s					
		revision clause	2	+;||	s				
		terminology	3			s			
COD98252	Taking up, the pursuit and the supervision of the business of electronic money institutions	strength of regulation	1	s					
		derogations	2	+;||	s				
COD98300	EC/Turkey relations: implementation of measures to promote economic and social development	amount of money	1	s					
		minority rights	2	+;||	s				
		nuclear strategy	3	+;||		s			
COD99204	System for the identification and registration of bovine animals	timing for the introduction of the first compulsory step	1	s					
		timing for the introduction of the second compulsory step	2		s				
		level of detail on the label during the first stage of the system	3	-;||		s			
		level of detail on the label during the second stage of the system	4		-;||		s		
COD99252	Interoperability of the trans-European conventional rail system	scope of harmonization	1	s					
		timing of the implementation of the Directive	2	+;||	s				
		discretion available to the TSI committee	3			s			

A.6 The magnitude of nonseparability at the proposal level

Figure A.3: Mean average error at different levels of nonseparability and model's predictive accuracy at the proposal level

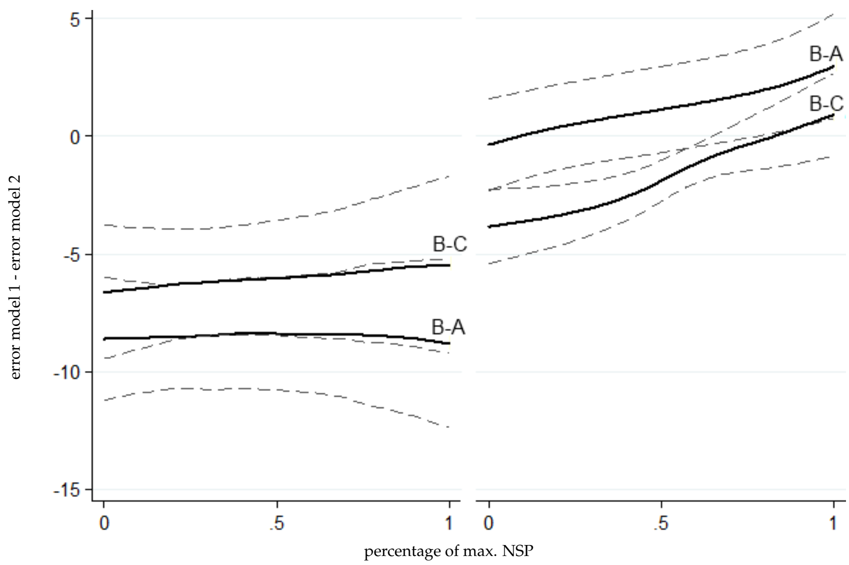
EXPLANATORY NOTE

The figure depicts each model's mean average error per proposal at different levels of nonseparability and 90% CI. The graph is the result of a local polynomial smoother which uses an Epanechnikov kernel (Epanechnikov, 1969) with a degree of 0. It is best interpreted as a moving average through the simulated data. The 41 issues categorized with reciprocal NSP are depicted in the right and the 40 issues categorized with non-reciprocal NSP in the left figure.



EXPLANATORY NOTE

The figure is the result of a local polynomial smoother which uses an Epanechnikov kernel (Epanechnikov, 1969) with a degree of 0. It includes 90% CI and is best interpreted as a moving average through the simulated data. The 41 issues categorized with reciprocal NSP are depicted in the right and the 40 issues categorized with non-reciprocal NSP in the left figure.



LABELS

A = agenda-setting model; B = unconstrained bargaining model; C = constrained bargaining model.

A.7 Experimental instructions

The instructions for the participants in my laboratory experiment consisted of two parts. The first part introduced the subjects to the general environment, such as that they are taking part in an lab experiment, that they will face a collective decision-making, that they will receive a payment for the points earned in the experiment after they have been multiplied with an exchange rate, that they are already entitled to a show-up fee, etc. In particular they were introduced to the payoff table and its significance.

The second part described the rules of the corresponding treatment as well as the screens they would encounter when they are asked to make their decision. For that purpose, the introductions included screenshots of the upcoming experiment.

Subjects were only given rules for one treatment at a time, the one design that was applied subsequently. Nevertheless, subjects were told at the beginning of the experiment that the decision-making rules would be altered during the experiment and that they would receive the corresponding new instructions when this would happen.

The following pages show the instructions as they were handed out to the participants. As the experimental sessions were conducted in Heidelberg the instructions are only available in German.

After distributing the instructions they were also read aloud by a supervisor. Questions were answered only privately. After that, the subjects went through the rules on their screen again and had to pass some questions testing their understanding of the upcoming game. The experiment did not start before all subjects had passed this test. During the whole experiment communication between subjects was forbidden.

PART I

- general setup p. 315-317

PART II

- pooling p. 318-319
- simultaneous delegation p. 320-322
- sequential delegation p. 323-325

Figure A.4: General instructions

Erläuterungen

Herzlich willkommen zum heutigen Experiment! Falls Sie im Folgenden etwas nicht verstehen, eine Fragen haben oder ein Problem auftaucht machen Sie bitte per Handzeichen auf sich aufmerksam. Wir werden dann zu Ihnen kommen. Fragen bitte nicht laut stellen! Während des Experiments ist keine Kommunikation mit anderen Teilnehmern erlaubt. Bitte stellen sie Ihre Handys lautlos oder schalten Sie sie aus.

In unserem Experiment können Sie Geld verdienen. Wie viel Sie erhalten, hängt von Ihren Entscheidungen im Experiment ab.

AUSZAHLUNGEN

- Sie erhalten in jedem Fall eine "show-up-fee", für Ihr Erscheinen, in Höhe von 4 €.
- Für das Ausfüllen eines Fragebogens erhalten Sie weitere 2 €.
- Zusätzlich können Sie in den einzelnen Runden des Experiments Punkte sammeln.
  - Am Ende des Experiments werden zufällig drei Runden ausgewählt. Die in diesen Runden gesammelten Punkte werden Ihnen gutgeschrieben.
  - Die Punkte werden im Verhältnis 5:1 in Euro umgerechnet. Für jeweils 5 gesammelte Punkte erhalten sie also 1 € zusätzlich.
- Das Geld wird am Ende des Experiments bar ausgezahlt. Jeder Teilnehmer erfährt nur seinen eigenen Auszahlungsbetrag.

ABLAUF

- Das Experiment ist nicht auf eine feste Anzahl Runden beschränkt.
- Das Experiment besteht aus 2 Phasen. Zuerst stimme Sie in einer Gruppe mit 6 Mitgliedern über mehrere Alternativen ab. Die Gruppe muss dabei kollektiv ein Entscheidungsproblem lösen.
- Nach einigen Runden werden die Abstimmungsregeln verändert. Sie werden dann noch einmal auf die neuen Regeln hingewiesen.
- Anschließend füllen Sie bitte noch einen Fragebogen aus. Darin werden wir Ihnen Fragen zum heutigen Experiment und zu einem aktuellem politischen Problem stellen.
- Es ist sehr wichtig, dass alle Teilnehmer die Regeln verstehen. Deshalb werden wir diese ausführlich erklären. Sie erhalten hierdurch aber keine geringeren Auszahlungen.

- 1 -

DIE ALTERNATIVEN

In jeder Runde müssen Sie zwischen verschiedenen Alternativen wählen. Diese werden Ihnen am Bildschirm wie in Abbildung 1 (hier rechts) angezeigt.

Abbildung 1: Eine Alternative.

	Punkteverteilung
Spieler 1	7
Spieler 2	28
Spieler 3	35
Spieler 4	22
Spieler 5	8
Spieler 6	18

Jede Alternative weist jedem der 6 Spieler eindeutig eine bestimmte Punkteanzahl zu. Die für den Spieler relevante Punkteanzahl steht immer in derselben Zeile wie seine Spieler-Nummer.
Im Beispiel: *Spieler 1* erhält 7 Punkte, *Spieler 2* erhält 28 Punkte, usw.

Insgesamt stehen 9 Alternativen zur Auswahl. Wie in der Abbildung 2 (hier rechts) zu sehen entstehen die 9 Alternativen als Kombination der **Spalten (Buchstaben A-C)** und **Zeilen (Ziffern 1-3)**. Somit sind A1, A2, A3, B1, B2, B3, C1, C2 und C3 mögliche Kombinationen.

Abbildung 2: Die mögliche Ergebnisse für Spieler 4.

		A	B	C
1	Player 1			
	Player 2			
	Player 3			
	Player 4	22	77	4
	Player 5			
	Player 6			
2	Player 1			
	Player 2			
	Player 3			
	Player 4	19	1	5
	Player 5			
	Player 6			
3	Player 1			
	Player 2			
	Player 3			
	Player 4	6	3	17
	Player 5			
	Player 6			

Um die Tabelle besser erklären zu können sind dort bisher nur die Punkte für *Spieler 4* eingetragen. Die für diesen Spieler relevante Punkteanzahl steht für jede Alternative in derselben Zeile wie die Spieler-Nummer.
Würde die Alternative B2 gewählt erhält der *Spieler 4* 1 Punkt. Falls Alternative C3 gewählt würde, bekäme *Spieler 4* 17 Punkte.

Mit ihrer Abstimmung beeinflussen die Spieler das Ergebnis aller Mitglieder der Gruppe. Für ihre eigene Punkteanzahl sind die Punkte der anderen Spieler aber nicht relevant!

Alle **9 Alternativen** bilden die komplette Auszahlungstabelle. In der vollständigen Auszahlungstabelle ist in jeder Alternative für jeden Spieler eine Punkteanzahl abgetragen. Auf der nächsten Seite sehen Sie ein Beispiel.

Die Auszahlungstabelle

Die Auszahlungstabelle bildet die Grundlage für alle im Experiment getroffenen Entscheidungen. Um diesen wichtigen Aspekt zu verdeutlichen hier ein Beispiel zu Abbildung 3:
Wird die Alternative B1 gewählt, erhält *Spieler 1* eine Punkteanzahl von 2, *Spieler 2* eine Punkteanzahl von 8, *Spieler 3* eine Punkteanzahl von 12, usw. Angenommen Sie sind *Spieler 4*, dann würden Sie eine Punkteanzahl in Höhe von 77 erhalten.

Bitte beachten Sie!

Hat die Gruppe eine Alternative gewählt, erhält jeder Spieler die dort für ihn festgelegte Punkteanzahl. Die Punkte, die einem Spieler gutgeschrieben werden, sind unabhängig davon, ob er auch für diese Alternative gestimmt hat. Entscheidend ist, dass die Gruppe diese Alternative gewählt hat.

In unserem Experiment werden Sie in einer Gruppe abstimmen. Bis sich die anderen Teilnehmer entschieden haben kann es zu Wartezeiten kommen.
Wir bemühen uns, diese so kurz wie möglich zu halten. Leider können wir diese aber nicht vermeiden. Bitte warten Sie in diesem Fall, bis sich alle Teilnehmer entschieden haben.

Abbildung 3: Die komplette Auszahlungstabelle

		A	**B**	**C**
1	Player 1	7	2	5
	Player 2	28	8	41
	Player 3	35	12	24
	Player 4	22	77	4
	Player 5	8	1	26
	Player 6	18	9	20
2	Player 1	58	45	10
	Player 2	2	1	3
	Player 3	17	8	70
	Player 4	19	1	5
	Player 5	4	45	12
	Player 6	11	10	25
3	Player 1	8	21	18
	Player 2	10	49	23
	Player 3	66	21	22
	Player 4	6	3	17
	Player 5	10	6	20
	Player 6	24	17	15

- 3 -

Figure A.5: Pooling instructions

DIE GRUPPEN

- Es werden Gruppen mit je 6 Spielern gebildet.
- Die Mitglieder einer Gruppe werden mit *Spieler 1*, *Spieler 2*, *Spieler 3*, *Spieler 4*, *Spieler 5* und *Spieler 6* bezeichnet.
- Jedes Gruppenmitglied verfügt über genau eine Stimme.
- Zu Beginn jeder Runde werden die Gruppen zufällig neu zusammengestellt. Sie kennen die Identität der anderen Mitglieder Ihrer Gruppe nicht.
- Am Ende einer Runde erfahren Sie nur das Ergebnis Ihrer Gruppe, aber nicht die Ergebnisse der anderen Gruppen.

DIE REGELN

- Alle Spieler stimmen über alle 9 Alternativen ab.
- Das Endergebnis der Runde ist die Alternative, die von einer absoluten Mehrheit - also mindestens 4 Spielern - gewählt wird.
- Alle Spieler kennen alle möglichen Auszahlungen. Sie wissen also, wie viele Punkte Sie und jeder andere Spieler bei der Wahl einer Alternative erhalten.
- Im Verlauf einer Runde wird so oft über eine Auszahlungstabelle erneut abgestimmt, bis eine Alternative mindestens 4 Stimmen in einer Abstimmung erhält. Dann ist die Runde beendet und die Spieler erhalten ihre Punkte zugewiesen.
- In jeder neuen Runde stimmen die Spieler über eine neue Auszahlungstabelle ab.
- Alle Mitglieder einer Gruppe stimmen in jeder Abstimmungsrunde gleichzeitig ab.

WELCHE INFORMATIONEN HABE ICH FÜR MEINE ENTSCHEIDUNG?
Bevor die Teilnehmer abstimmen, sind sie über die Payoffs aller Mitglieder der Gruppe informiert.
Alle Mitglieder einer Gruppe stimmen in jeder Abstimmungsrunde gleichzeitig ab. Für welche Alternativen die anderen Spieler derselben Gruppe gestimmt haben, erfahren Sie jeweils erst, wenn alle Mitglieder Ihre Stimme abgegeben haben.

- 4 -

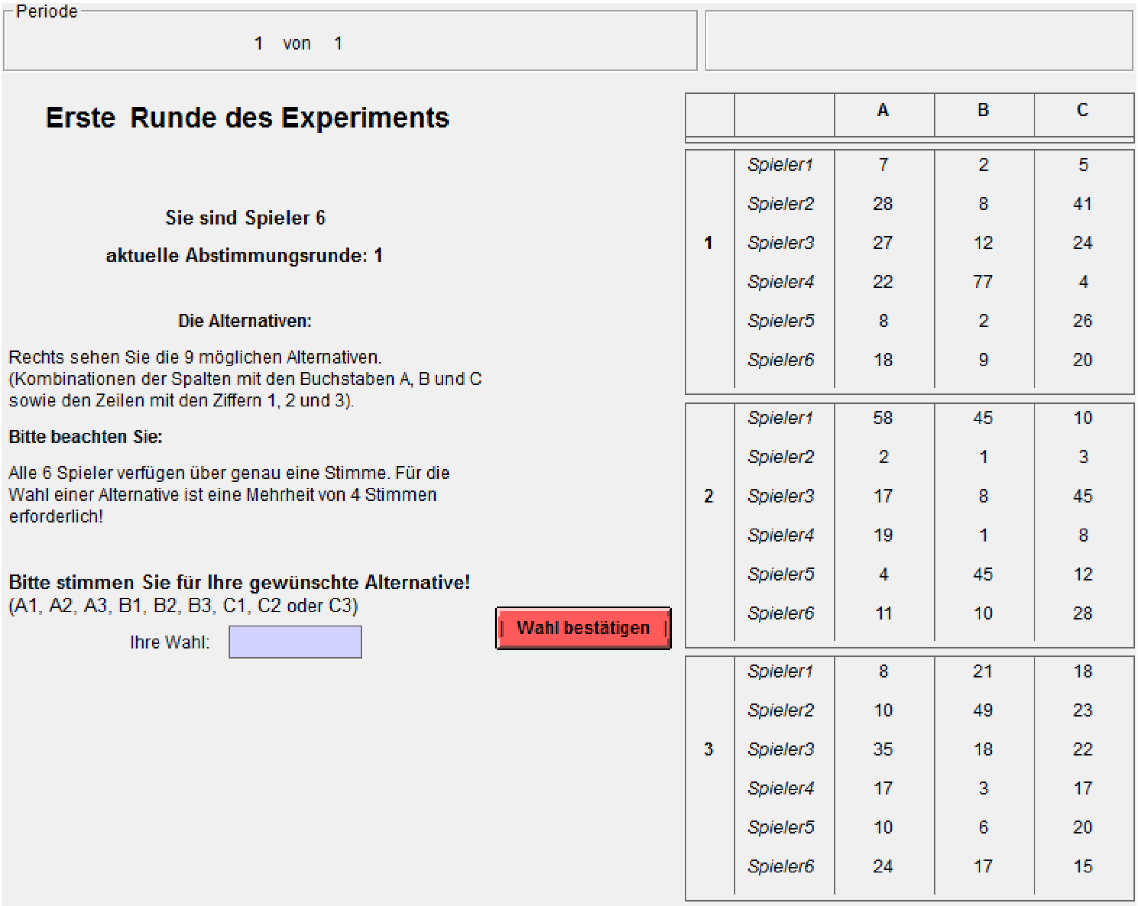
Die Entscheidungsfindung

Abbildung 4 zeigt den Entscheidungsbildschirm wie Sie ihn auch später an Ihrem PC sehen werden.
Da Ihre Entscheidung auf der Auszahlungstabelle beruht, wird diese für Sie immer eingeblendet. Links davon sehen Sie Ihre Spielernummer und die Nummer der aktuellen Abstimmungsrunde.
In das freie Textfeld können Sie ihre gewählte Alternative eingeben (Bitte achten Sie bei den Buchstaben auf Großschreibung). Haben alle Mitglieder der Gruppe ihre Wahl getroffen, wird die Abstimmung ausgewertet. Sie erhalten dann

- entweder die Meldung „Ihre Gruppe hat sich entschieden!“ und Ihnen wird das Ergebnis mitgeteilt. Die Spieler erhalten dann die entsprechende Punkteanzahl gutgeschrieben.

- oder die Meldung „Keine der Alternativen hat die erforderliche Mehrheit von mindestens 4 Stimmen erhalten!“ mit der Aufforderung erneut abzustimmen.

Die Abstimmung wird so lange fortgesetzt, bis eine Alternative die nötige Mehrheit von mindestens 4 Stimmen erreicht. Sind mehrere Abstimmungen nötig, werden links unten die Ergebnisse früherer Abstimmungen eingeblendet.



		A	B	C
1	Spieler1	7	2	5
	Spieler2	28	8	41
	Spieler3	27	12	24
	Spieler4	22	77	4
	Spieler5	8	2	26
	Spieler6	18	9	20
2	Spieler1	58	45	10
	Spieler2	2	1	3
	Spieler3	17	8	45
	Spieler4	19	1	8
	Spieler5	4	45	12
	Spieler6	11	10	28
3	Spieler1	8	21	18
	Spieler2	10	49	23
	Spieler3	35	18	22
	Spieler4	17	3	17
	Spieler5	10	6	20
	Spieler6	24	17	15

Abbildung 4: Entscheidungsbildschirm

- 5 -

Figure A.6: Simultaneous delegation instructions

DIE GRUPPEN

- Es werden Gruppen mit je 6 Mitgliedern gebildet. Jedes Gruppenmitglied verfügt über genau eine Stimme.
- Die Mitglieder einer Gruppe werden mit *Spieler 1*, *Spieler 2*, *Spieler 3*, *Spieler 4*, *Spieler 5* und *Spieler 6* bezeichnet.
- Jede Gruppe wird in zwei Teams aufgeteilt. Jedes Team hat 3 Mitglieder.
- Die Teams haben unterschiedlich Aufgaben. Jedes Team ist für einen Teil der Entscheidungsfindung verantwortlich.
  - Ein Team muss sich auf die zu wählende Spalte (Buchstaben A - C) verständigen.
  - Das andere Team muss sich auf die zu wählende Zeile (Ziffer 1 -3) verständigen.

DIE REGELN

- Die Spieler können jeweils nur eine der beiden Entscheidungen selbst treffen!
- Es wird die Zeile / Spalte gewählt, die von einer absoluten Mehrheit eines Teams - also mindestens 2 von 3 Spielern - gewählt wird.
- Im Verlauf einer Runde wird so oft erneut abgestimmt, bis eine Zeile / Spalte mindestens diese 2 Stimmen erhält. Dann ist die Abstimmung des Teams beendet.
- Am Ende einer Runde erfahren Sie nur das Ergebnis Ihrer Gruppe, aber nicht die Ergebnisse der anderen Gruppen.
- Zu Beginn jeder Runde werden die Gruppen und die Teams zufällig neu zusammengestellt.
- In jeder neuen Runde stimmen die Spieler über eine neue Auszahlungstabelle ab.

WIE WIRD DAS ENDERGEBNIS ERMITTELT?

Wenn Ihr Team sich geeinigt hat, ist der erste Teil der Entscheidungsfindung abgeschlossen. Für die Bestimmung des Endergebnisses muss sich nun noch das andere Team entscheiden. Sobald sich beide Teams entschieden haben, wird das Endergebnis bestimmt und den Spielern mitgeteilt.

- 4 -

DIE ENTSCHEIDUNGSFINDUNG

Abbildung 4 zeigt den Entscheidungsbildschirm wie Sie ihn auch später an Ihrem PC sehen werden.

Da Ihre Entscheidung auf der Auszahlungstabelle beruht, wird diese für Sie immer eingeblendet. Links davon sehen Sie Ihre Spielernummer und die Nummer der aktuellen Abstimmungsrunde. Unter Ihrer Spielernummer sehen sie die Angabe, ob Sie über die Zeilen oder die Spalten abstimmen werden. Zudem werden auch die beiden anderen Mitglieder Ihres Teams genannt.

In das freie Textfeld können Sie ihre gewählte Alternative eingeben (Bitte achten Sie bei den Buchstaben auf Großschreibung). Wie in Abbildung 4 unten zu sehen ist, dürfen Sie hier nur die Zeile (1, 2 oder 3) wählen. Das andere Team entscheidet über die Spalte. Haben alle Mitglieder Ihres Teams ihre Wahl getroffen, wird die Abstimmung ausgewertet. Sie erhalten dann entweder

- die Meldung „Ihr Team hat sich entschieden!" und Ihnen wird das Ergebnis mitgeteilt oder
- die Meldung „Keine der Alternativen hat die erforderliche Mehrheit von mindestens 2 Stimmen erhalten!" mit der Aufforderung erneut abzustimmen.

Die Abstimmung wird so lange fortgesetzt, bis eine Alternative die nötige Mehrheit von mindestens 2 Stimmen erreicht. Sind mehrere Abstimmungen nötig, werden links unten die Ergebnisse früherer Abstimmungen eingeblendet.

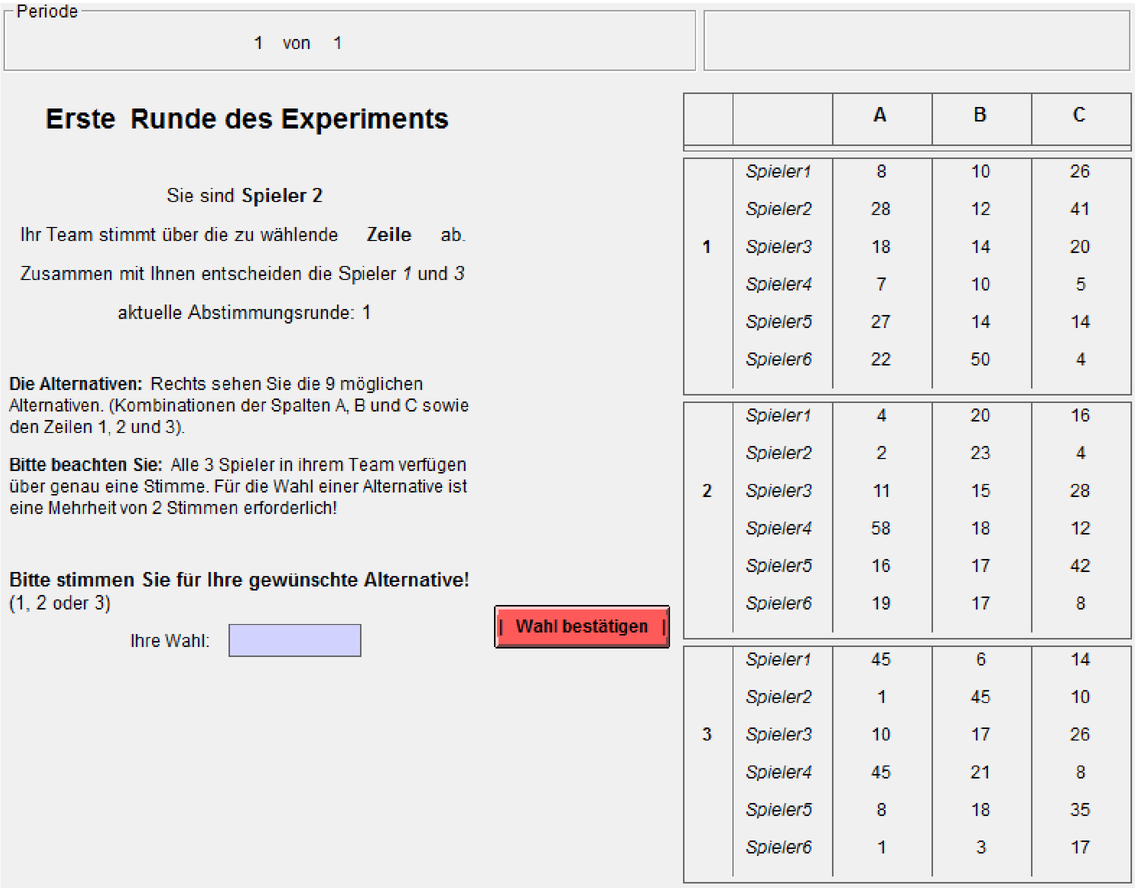


Abbildung 4: Entscheidungsbildschirm

Welche Informationen habe ich für meine Entscheidung?
Bevor die Teilnehmer abstimmen, sind sie über die Payoffs aller Mitglieder des eigenen Teams und des anderen Teams informiert. Dabei wissen die Spieler auch, welche beiden anderen Spieler im selben Team sind. Die übrigen drei Spieler der Gruppe bilden dann das andere Team.
Alle Mitglieder eines Teams stimmen in jeder Abstimmungsrunde gleichzeitig ab. Für welche Alternativen die anderen Spieler desselben Teams gestimmt haben, erfahren Sie jeweils erst, wenn alle 3 Teammitglieder Ihre Stimme abgegeben haben. Für welche Alternativen das andere Team gestimmt hat, erfahren Sie erst, wenn sich beide Teams geeinigt haben.

Beispiel zu Abbildung 4

- Das Endergebnis hängt von den Abstimmungen in beiden Teams ab!
- Angenommen *Team 1* wählt Zeile 2 und *Team 2* wählt Spalte A. Dann ist Alternative A2 gewählt. I n diesem Fall erhält *Spieler 1* 4 Punkte, *Spieler 2* 2 Punkte, *Spieler 3* 11 Punkte, usw.
- Wählt aber Team 2 statt Spalte A die Spalte B erhält *Spieler 1* 20 Punkte, *Spieler 2* 23 Punkt, *Spieler 3* 15 Punkte, usw.
- Die Punkteanzahl ändert sich also für alle Spieler, auch wenn nur ein Team seine Entscheidung verändert.

Bitte beachten Sie!
Hat das Team eine Zeile / Spalte gewählt, ist diese für alle Mitglieder festgelegt. Dies ist unabhängig davon, ob jedes Mitglied des Teams auch für diese Zeile / Spalte gestimmt hat. Entscheidend ist, dass das Team diese Zeile / Spalte mit min. 2 Stimmen gewählt hat.

Figure A.7: Sequential delegation instructions

DIE GRUPPEN

- Es werden Gruppen mit je 6 Mitgliedern gebildet. Jedes Gruppenmitglied verfügt über genau eine Stimme.
- Die Mitglieder einer Gruppe werden mit *Spieler 1*, *Spieler 2*, *Spieler 3*, *Spieler 4*, *Spieler 5* und *Spieler 6* bezeichnet.
- Jede Gruppe wird in zwei Teams aufgeteilt. Jedes Team hat 3 Mitglieder.
- Die Teams haben unterschiedlich Aufgaben. Jedes Team ist für einen Teil der Entscheidungsfindung verantwortlich.
  - Ein Team muss sich auf die zu wählende Spalte (Buchstaben A - C) verständigen.
  - Das andere Team muss sich auf die zu wählende Zeile (Ziffer 1 -3) verständigen.
- Zuerst entscheidet das Team, welches über die zu wählende Spalte abstimmt.
- Das Ergebnis dieser Abstimmung wird allen 6 Mitgliedern der Gruppe (also beiden Teams) mitgeteilt.
- Erst danach stimmt das andere Team über die Zeile ab.

DIE REGELN

- Die Spieler können jeweils nur eine der beiden Entscheidungen selbst treffen!
- Es wird die Zeile / Spalte gewählt, die von einer absoluten Mehrheit eines Teams - also mindestens 2 von 3 Spielern - gewählt wird.
- Im Verlauf einer Runde wird so oft erneut abgestimmt, bis eine Zeile / Spalte mindestens diese 2 Stimmen erhält. Dann ist die Abstimmung des Teams beendet.
- Am Ende einer Runde erfahren Sie nur das Ergebnis Ihrer Gruppe, aber nicht die Ergebnisse der anderen Gruppen.
- Zu Beginn jeder Runde werden die Gruppen und die Teams zufällig neu zusammengestellt.
- In jeder neuen Runde stimmen die Spieler über eine neue Auszahlungstabelle ab.

WIE WIRD DAS ENDERGEBNIS ERMITTELT?

Wenn das erste Team sich für eine Spalte entschieden hat, ist der erste Teil der Entscheidungsfindung abgeschlossen. Für die Bestimmung des Endergebnisses muss sich nun noch das zweite Team für eine Zeile entscheiden.

- 4 -

DIE ENTSCHEIDUNGSFINDUNG

Abbildung 4 zeigt den Entscheidungsbildschirm wie Sie ihn auch später an Ihrem PC sehen werden.

Da die Teams nicht mehr gleichzeitig abstimmen wird Ihnen mitgeteilt, ob ihr Team zuerst abstimmt oder das andere Team sich bereits entschieden hat. Dann sehen Sie deren Entscheidung. In Abbildung 4 hat das andere Team sich bereits für Spalte B entschieden. Sie stimmen nun noch über die zu wählende Zeile ab. Haben alle Mitglieder der Gruppe ihre Wahl getroffen, wird die Abstimmung ausgewertet. Sie erhalten dann

- entweder die Meldung „Ihre Gruppe hat sich entschieden!“ und Ihnen wird das Ergebnis mitgeteilt. Die Spieler erhalten dann die entsprechende Punkteanzahl gutgeschrieben.
- oder die Meldung „Keine der Alternativen hat die erforderliche Mehrheit von mindestens 2 Stimmen erhalten!“ mit der Aufforderung erneut abzustimmen.

Die Abstimmung wird so lange fortgesetzt, bis eine Alternative die nötige Mehrheit von mindestens 2 Stimmen erreicht. Sind mehrere Abstimmungen nötig, werden links unten die Ergebnisse früherer Abstimmungen eingeblendet. Die Runden die das andere Team zur Einigung benötigt hat werden leer angezeigt, da Ihr Team hier noch keine Wahl getroffen hat.

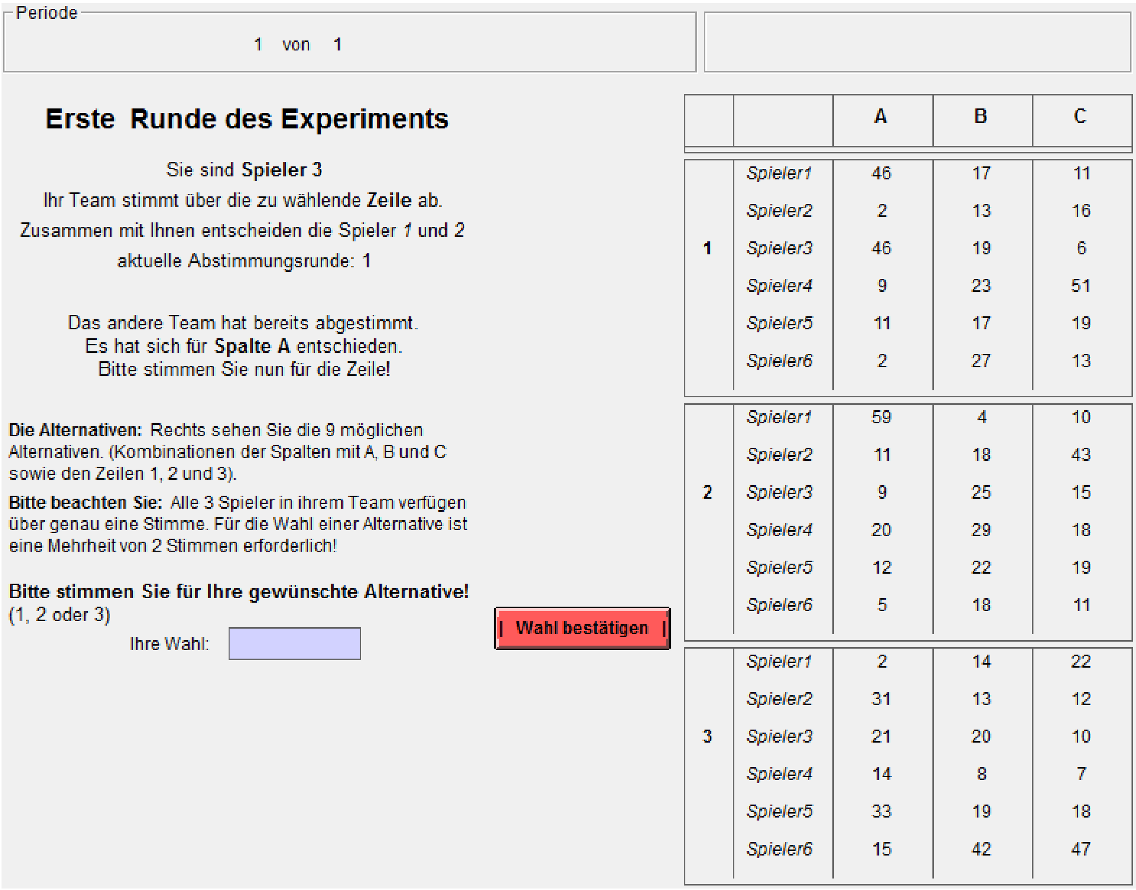


Abbildung 4: Entscheidungsbildschirm

Welche Informationen habe ich für meine Entscheidung?
Bevor die Teilnehmer abstimmen, sind sie über die Payoffs aller Mitglieder des eigenen Teams und des anderen Teams informiert. Dabei wissen die Spieler auch, welche beiden anderen Spieler im selben Team sind. Die übrigen drei Spieler der Gruppe bilden dann das andere Team.
Alle Mitglieder eines Teams stimmen in jeder Abstimmungsrunde gleichzeitig ab. Für welche Alternativen die anderen Spieler desselben Teams gestimmt haben, erfahren Sie jeweils erst, wenn alle 3 Teammitglieder Ihre Stimme abgegeben haben.

Beispiel zu Abbildung 4

- Das Endergebnis hängt von den Abstimmungen in beiden Teams ab!
- Angenommen *Team 1* hat sich bereits für Spalte A entschieden und ihr *Team* wählt nun Spalte 2. Dann ist Alternative A2 gewählt. In diesem Fall erhält *Spieler 1* 59 Punkte, *Spieler 2* 11 Punkte, *Spieler 3* 9 Punkte, usw.
- Wählt aber ihr *Team* statt Zeile 2 die Zeile 3 ist Alternative A3 gewählt und *Spieler 1* erhält 2 Punkte, *Spieler 2* erhält 31 Punkt, *Spieler 3* erhält 21 Punkte, usw.
- Die Punkteanzahl ändert sich also für alle Spieler, auch wenn nur ein Team seine Entscheidung verändert.

Bitte beachten Sie!
Hat das Team eine Zeile / Spalte gewählt, ist diese für alle Mitglieder festgelegt. Dies ist unabhängig davon, ob jedes Mitglied des Teams auch für diese Zeile / Spalte gestimmt hat. Entscheidend ist, dass das Team diese Zeile / Spalte mit min. 2 Stimmen gewählt hat.

A.8 All payoff tables used in the experiment

EXPLANATORY NOTE

Below each table it is indicated if the payoff table constitutes a constant-sum or non-constant-sum game (cf. Sec. 4.2.4) and if a core alternative for the different procedures exists (cf. Sec. 5.2).

Table A.5: Payoff tables used in the experiment

		A	B	C
1	Player 1	7	2	1
	Player 2	28	8	41
	Player 3	27	12	24
	Player 4	22	44	4
	Player 5	4	2	26
	Player 6	6	9	20
2	Player 1	58	35	5
	Player 2	2	11	3
	Player 3	10	4	45
	Player 4	19	1	8
	Player 5	8	45	12
	Player 6	11	10	28
3	Player 1	8	21	18
	Player 2	10	32	23
	Player 3	10	18	15
	Player 4	17	13	11
	Player 5	10	6	20
	Player 6	32	17	4

Payoff table 1
Non-constant-sum table
Core: -

		A	B	C
1	Player 1	9	4	3
	Player 2	30	10	43
	Player 3	29	14	26
	Player 4	24	46	6
	Player 5	6	4	28
	Player 6	8	11	22
2	Player 1	60	37	7
	Player 2	4	13	5
	Player 3	12	6	47
	Player 4	21	3	10
	Player 5	10	47	14
	Player 6	13	12	30
3	Player 1	10	23	20
	Player 2	12	34	25
	Player 3	12	20	17
	Player 4	19	15	13
	Player 5	12	8	22
	Player 6	34	19	6

Payoff table 2
Non-constant-sum table
Core: -

		A	B	C
1	Player 1	10	30	20
	Player 2	6	4	13
	Player 3	35	3	10
	Player 4	17	14	16
	Player 5	20	33	17
	Player 6	8	47	19
2	Player 1	28	29	22
	Player 2	13	6	34
	Player 3	20	12	30
	Player 4	9	29	19
	Player 5	5	3	12
	Player 6	34	2	9
3	Player 1	16	13	15
	Player 2	19	32	16
	Player 3	7	46	18
	Player 4	27	28	21
	Player 5	12	5	33
	Player 6	19	11	29

Payoff table 3
Non-constant-sum table
Core: SEQ {A1}

		A	B	C
1	Player 1	9	29	19
	Player 2	5	3	12
	Player 3	34	2	9
	Player 4	16	13	15
	Player 5	19	32	16
	Player 6	7	46	18
2	Player 1	27	28	21
	Player 2	12	5	33
	Player 3	19	11	29
	Player 4	9	29	19
	Player 5	5	3	12
	Player 6	34	2	9
3	Player 1	16	13	15
	Player 2	19	32	16
	Player 3	7	46	18
	Player 4	27	28	21
	Player 5	12	5	33
	Player 6	19	11	29

Payoff table 4
Non-constant-sum table
Core: SEQ {A1}

		A	B	C
1	Player 1	3	10	26
	Player 2	28	12	41
	Player 3	18	14	20
	Player 4	7	10	5
	Player 5	27	14	14
	Player 6	27	50	4
2	Player 1	11	20	16
	Player 2	6	23	8
	Player 3	11	15	28
	Player 4	44	18	12
	Player 5	13	17	32
	Player 6	25	17	14
3	Player 1	45	6	14
	Player 2	5	45	10
	Player 3	10	17	26
	Player 4	40	21	8
	Player 5	8	18	35
	Player 6	2	3	17

Payoff table 5
Constant-sum table
Core: SEQ {A1}

		A	B	C
1	Player 1	8	15	31
	Player 2	33	17	46
	Player 3	23	19	25
	Player 4	12	15	10
	Player 5	32	19	19
	Player 6	32	55	9
1	Player 1	16	25	21
	Player 2	11	28	13
	Player 3	16	20	33
	Player 4	49	23	17
	Player 5	18	22	37
	Player 6	30	22	19
1	Player 1	50	11	19
	Player 2	10	50	15
	Player 3	15	22	31
	Player 4	45	26	13
	Player 5	13	23	40
	Player 6	7	8	22

Payoff table 6
Constant-sum table
Core: SEQ {A1}

		A	B	C
1	Player 1	1	43	2
	Player 2	7	11	12
	Player 3	45	9	22
	Player 4	10	17	9
	Player 5	39	21	46
	Player 6	8	9	19
2	Player 1	10	14	41
	Player 2	42	34	10
	Player 3	14	16	19
	Player 4	18	32	20
	Player 5	9	1	13
	Player 6	17	13	7
3	Player 1	5	38	12
	Player 2	25	12	22
	Player 3	8	18	5
	Player 4	25	4	18
	Player 5	3	16	3
	Player 6	44	22	50

Payoff table 7
Constant-sum table
Core: -

		A	B	C
1	Player 1	3	45	4
	Player 2	9	13	14
	Player 3	47	11	24
	Player 4	12	19	11
	Player 5	41	23	48
	Player 6	10	11	21
1	Player 1	12	16	43
	Player 2	44	36	12
	Player 3	16	18	21
	Player 4	20	34	22
	Player 5	11	3	15
	Player 6	19	15	9
1	Player 1	7	40	14
	Player 2	27	14	24
	Player 3	10	20	7
	Player 4	27	6	20
	Player 5	5	18	5
	Player 6	46	24	52

Payoff table 8
Constant-sum table
Core: -

		A	B	C
1	Player 1	48	19	13
	Player 2	4	15	18
	Player 3	48	21	8
	Player 4	11	25	53
	Player 5	13	19	21
	Player 6	4	29	15
2	Player 1	61	6	12
	Player 2	13	20	45
	Player 3	11	27	17
	Player 4	22	31	20
	Player 5	14	24	21
	Player 6	7	20	13
3	Player 1	4	16	24
	Player 2	33	15	14
	Player 3	23	22	12
	Player 4	16	10	9
	Player 5	35	21	20
	Player 6	17	44	49

Payoff table 9
Constant-sum table
Core: POL, SIM and SEQ {B2}

		A	B	C
1	Player 1	46	17	11
	Player 2	2	13	16
	Player 3	46	19	6
	Player 4	9	23	51
	Player 5	11	17	19
	Player 6	2	27	13
1	Player 1	59	4	10
	Player 2	11	18	43
	Player 3	9	25	15
	Player 4	20	29	18
	Player 5	12	22	19
	Player 6	5	18	11
1	Player 1	2	14	22
	Player 2	31	13	12
	Player 3	21	20	10
	Player 4	14	8	7
	Player 5	33	19	18
	Player 6	15	42	47

Payoff table 10
Constant-sum table
Core: POL, SIM and SEQ {B2}

		A	B	C
1	Player 1	2	33	4
	Player 2	11	11	12
	Player 3	37	9	24
	Player 4	25	17	28
	Player 5	30	31	28
	Player 6	5	9	14
2	Player 1	31	14	35
	Player 2	29	33	10
	Player 3	21	20	20
	Player 4	6	25	4
	Player 5	1	2	17
	Player 6	22	16	24
3	Player 1	3	38	22
	Player 2	13	12	12
	Player 3	28	13	15
	Player 4	39	8	18
	Player 5	3	19	30
	Player 6	24	20	13

Payoff table 11
Constant-sum table
Core: POL {A3}, SIM and SEQ {C1}

		A	B	C
1	Player 1	7	38	9
	Player 2	16	16	17
	Player 3	42	14	29
	Player 4	30	22	33
	Player 5	35	36	33
	Player 6	10	14	19
1	Player 1	36	19	40
	Player 2	34	38	15
	Player 3	26	25	25
	Player 4	11	30	9
	Player 5	6	7	22
	Player 6	27	21	29
1	Player 1	8	43	27
	Player 2	18	17	17
	Player 3	33	18	20
	Player 4	44	13	23
	Player 5	8	24	35
	Player 6	29	25	18

Payoff table 12
Constant-sum table
Core: POL {A3}, SIM and SEQ {C1}

		A	B	C
1	Player 1	22	9	14
	Player 2	16	11	10
	Player 3	30	6	14
	Player 4	17	2	42
	Player 5	3	21	9
	Player 6	22	17	18
2	Player 1	5	20	33
	Player 2	12	41	1
	Player 3	33	7	8
	Player 4	15	14	1
	Player 5	14	22	38
	Player 6	19	18	10
3	Player 1	8	7	20
	Player 2	4	26	14
	Player 3	19	17	13
	Player 4	10	12	30
	Player 5	42	19	1
	Player 6	20	16	32

Payoff table 13
Non-constant-sum table
Core: POL {A1}, SIM and SEQ {C3}

		A	B	C
1	Player 1	25	12	17
	Player 2	19	14	13
	Player 3	33	9	16
	Player 4	20	5	45
	Player 5	6	24	12
	Player 6	25	20	21
2	Player 1	8	23	36
	Player 2	15	44	4
	Player 3	36	10	11
	Player 4	18	17	4
	Player 5	17	25	41
	Player 6	21	21	13
3	Player 1	11	10	23
	Player 2	7	29	17
	Player 3	22	20	16
	Player 4	13	15	33
	Player 5	45	22	4
	Player 6	23	19	35

Payoff table 14
Non-constant-sum table
Core: POL {A1}, SIM and SEQ {C3}

		A	B	C
1	Player 1	15	16	5
	Player 2	4	15	25
	Player 3	11	19	9
	Player 4	35	15	17
	Player 5	22	24	28
	Player 6	2	8	8
2	Player 1	6	5	18
	Player 2	32	11	16
	Player 3	6	34	13
	Player 4	9	11	18
	Player 5	3	15	4
	Player 6	35	16	17
3	Player 1	10	4	35
	Player 2	2	44	11
	Player 3	12	7	10
	Player 4	25	8	12
	Player 5	27	22	15
	Player 6	16	6	11

Payoff table 15
Non-constant-sum table
Core: POL and SIM {C2}

		A	B	C
1	Player 1	17	18	7
	Player 2	6	17	27
	Player 3	13	21	11
	Player 4	37	17	19
	Player 5	24	26	30
	Player 6	4	10	10
2	Player 1	8	7	20
	Player 2	34	13	18
	Player 3	8	36	15
	Player 4	11	13	20
	Player 5	5	17	6
	Player 6	37	18	19
3	Player 1	12	6	37
	Player 2	4	46	13
	Player 3	14	9	12
	Player 4	27	10	14
	Player 5	29	24	17
	Player 6	18	8	13

Payoff table 16
Non-constant-sum table
Core: POL and SIM {C2}

		A	B	C
1	Player 1	25	14	22
	Player 2	4	24	33
	Player 3	36	17	31
	Player 4	14	16	15
	Player 5	21	27	17
	Player 6	31	20	20
2	Player 1	11	27	21
	Player 2	33	36	5
	Player 3	4	21	27
	Player 4	15	11	35
	Player 5	8	15	21
	Player 6	33	10	25
3	Player 1	14	33	10
	Player 2	6	8	25
	Player 3	37	5	8
	Player 4	39	17	16
	Player 5	22	31	25
	Player 6	27	36	20

Payoff table 17
Non-constant-sum table
Core: POL {B3}, SEQ {A3}

		A	B	C
1	Player 1	23	12	20
	Player 2	2	22	31
	Player 3	34	15	29
	Player 4	12	14	13
	Player 5	19	25	15
	Player 6	29	18	18
2	Player 1	9	25	19
	Player 2	31	34	3
	Player 3	2	19	25
	Player 4	13	9	33
	Player 5	6	13	19
	Player 6	31	8	23
3	Player 1	12	31	8
	Player 2	4	6	23
	Player 3	35	3	6
	Player 4	37	15	14
	Player 5	20	29	23
	Player 6	25	34	18

Payoff table 18
Non-constant-sum table
Core: POL {B3}, SEQ {A3}

		A	B	C
1	Player 1	11	18	14
	Player 2	25	8	22
	Player 3	19	14	18
	Player 4	22	9	19
	Player 5	2	26	1
	Player 6	9	11	14
2	Player 1	19	10	20
	Player 2	10	9	12
	Player 3	14	22	20
	Player 4	11	24	2
	Player 5	18	3	5
	Player 6	12	18	24
3	Player 1	16	15	13
	Player 2	6	10	10
	Player 3	9	22	14
	Player 4	14	2	22
	Player 5	26	22	24
	Player 6	19	16	8

Payoff table 19
Non-constant-sum table
Core: POL {C2}, SEQ {B3}

		A	B	C
1	Player 1	14	21	17
	Player 2	28	11	25
	Player 3	22	17	21
	Player 4	25	12	22
	Player 5	5	29	4
	Player 6	12	14	17
2	Player 1	22	13	23
	Player 2	13	12	15
	Player 3	17	25	23
	Player 4	14	27	5
	Player 5	21	6	8
	Player 6	15	21	27
3	Player 1	19	18	16
	Player 2	9	13	13
	Player 3	12	25	17
	Player 4	17	5	25
	Player 5	29	25	27
	Player 6	22	19	11

Payoff table 20
Non-constant-sum table
Core: POL {C2}, SEQ {B3}

A.9 Payoff table characteristics

Table A.6: Payoff table characteristics

EXPLANATORY NOTE
The table describes the distribution of the payoff table characteristics across the decision-making procedures. For each procedure tables with and without a core alternative as well as constant-sum and non-constant-sum tables exist. Also, every combination of the two characteristics exists for every procedure. The core concept refers under the pooling procedure to the "standard" deterministic core. For simultaneous and sequential delegation it indicates the existence of the credible core. Due to these differences the table classification differs between procedures.

		CORE ALTERNATIVE			
		does exist		does not exist	
		CONSTANT-SUM TABLE		CONSTANT-SUM TABLE	
		yes	no	yes	no
PROCEDURE	pooling	9-12	13-20	5-8	1-4
	simultaneous delegation	9-12	13-16	5-8	1-4 and 17-20
	sequential delegation	5, 6, and 9-12	3, 4, 13, 14 and 17-20	7-8	1, 2, 15 and 16

A.10 Frequency of use of each payoff table

Table A.7: Frequency of use of each payoff table

EXPLANATORY NOTE
The table shows the number of observations for each payoff table. Each observation refers to one collective agreement on an alternative. The observations are separated according to the implemented decision-making procedure. Overall, I conducted 221 rounds of which 76 used pooling, 85 used simultaneous delegation and 60 used sequential delegation. Due to the random assignment rule the frequency of use differs and not all tables are used under all procedures. This does not constitute inevitably a problem as multiple tables share the same properties; e.g., if a tables contains a core alternative (cf. Tab. A.6). Sec. A.11 takes a closer look at the actual data structure.

		NUMBER OF PAYOFF TABLE																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ
PROCEDURE	pooling	4	0	3	2	4	0	5	0	7	2	7	2	4	4	6	4	7	4	5	6	76
	simultaneous delegation	0	10	0	3	0	8	2	3	2	9	0	11	2	9	0	9	2	7	0	8	85
	sequential delegation	0	4	0	2	2	2	0	2	6	2	4	4	4	4	2	6	2	6	4	4	60
	Σ	4	14	3	7	6	10	7	5	15	13	11	17	10	17	8	19	11	17	9	18	221

A.11 Data structure

In my 13 experimental sessions I collected, overall, 76 pooling, 85 simultaneous delegation and 60 sequential delegation observations. Despite the random assignment rule (cf. Sec. 4.2.4) I obtained observations for every possible combination of table characteristics. However, across all procedures the number for some specifications is rather small and it must be admitted that the distribution is skewed. Pooling and sequential delegation show a preponderance of core alternative tables and under simultaneous delegation more observations were collected as first decision rule. In my analysis of the experimental results in Chap. 5 and Chap. 6, I always state the respective number of observations to secure a valid investigation.

Table A.8: Data structure according to decision rule and payoff table properties

EXPLANATORY NOTE
The table below separates the 221 collected observations due to procedure, first or second decision rule, if the table contained an equilibrium alternative and if the table represented a constant-sum or non-constant-sum game.

		CORE-ALTERNATIVE			
		does exist		does not exist	
PROCEDURE	DECISION RULE	CONSTANT-SUM TABLE		CONSTANT-SUM TABLE	
		yes	no	yes	no
pooling	first	7	11	5	5
	second	11	29	4	4
	Σ	18	40	9	9
simultaneous delegation	first	15	13	11	21
	second	7	7	2	9
	Σ	22	20	13	30
sequential delegation	first	10	20	2	6
	second	8	10	2	2
	Σ	18	30	4	8

A.12 Trend analysis of the collective results

The following tables show the results of the trend analysis for the measures obtained in *Sec. 5.3.1* regarding decision-making efficiency, social welfare allocation, distribution of wealth and approval rate. As discussed in Sec. 5.4.1, I analyzed the development over the single rounds by using STATA's *somersd* package (implemented for STATA by Newson, 2002), a nonparametric rank-statistic. The package performs a maximum likelihood fit to obtain association measures.

Table A.9: Trend analysis

EXPLANATORY NOTE
In each table I list the Somers' D coefficient, its z-score and 95% CI. The observations are separated by procedure and for being the first or second decision rule of a session. Also, for approval rate the two delegation procedures show the results on the delegation level. Statistically significant (two-tailed) at the 0.1 level *, at the 0.05 level ** and at the 0.01 level ***.

Independent variable: NUMBER OF ROUND		Dependent variable: APPROVAL RATE				
		N	D	z-score	95% CI	
	pooling	76	0.00	0.74	-0.01	0.01
	first decision rule	28	0.00	0.74	-0.01	0.01
	second decision rule	48	-0.01	-0.83	-0.03	0.01
	simultaneous delegation	170	-0.01	-0.16	-0.07	0.06
PROCEDURE	first decision rule	120	-0.02	-0.46	-0.11	0.07
	second decision rule	50	-0.05	-0.76	-0.18	0.08
	sequential delegation	120	0.07*	1.66	-0.01	0.14
	first decision rule	76	0.05	1.08	-0.04	0.14
	second decision rule	44	-0.04	-0.42	-0.22	0.14

Independent variable: NUMBER OF ROUND		Dependent variable: DECISION-MAKING EFFICIENCY				
		N	D	z-score	95% CI	
PROCEDURE	all	221	0.16***	3.24	0.06	0.26
	first decision rule	126	0.01	0.11	-0.15	0.17
	second decision rule	95	-0.03	-0.30	-0.20	0.15
	pooling	76	0.15**	1.98	0.00	0.30
	first decision rule	28	0.08	0.64	-0.16	0.31
	second decision rule	48	0.20*	1.7	-0.03	0.44
	simultaneous delegation					
	delegation level	170	-0.4	-0.78	-0.14	0.07
	first decision rule	120	-0.05	-0.85	-0.15	0.06
	second decision rule	50	0.05	0.58	-0.13	0.23
	sequential delegation	60	0.09	1.06	-0.08	0.27
	first decision rule	38	0.05	0.38	-0.21	0.32
	second decision rule	22	-0.05	-0.26	-0.43	0.33
	delegation level	120	0.01	0.23	-0.18	0.16
	first decision rule	76	0.03	0.34	-0.13	0.18
	second decision rule	44	-0.05	-0.50	-0.27	0.16

Independent variable: NUMBER OF ROUND		Dependent variable: SOCIAL WELFARE ALLOCATION				
		N	D	z-score	95% CI	
PROCEDURE	pooling	49	-0.02	-0.19	-0.23	0.19
	first decision rule	16	-0.24	-1.14	-0.66	0.18
	second decision rule	33	0.17	1.24	-0.10	0.44
	simultaneous delegation					
	delegation level	100	-0.05	-0.60	-0.21	0.11
	first decision rule	68	-0.02	-0.22	-0.22	0.18
	second decision rule	32	-0.32	-0.26	-0.27	0.21
	sequential delegation	38	-0.16	-1.09	-0.44	0.13
	first decision rule	26	-0.10	-0.48	-0.51	0.31
	second decision rule	12	-0.33	-1.04	-0.96	0.29

Independent variable: NUMBER OF ROUND		Dependent variable: DISTRIBUTION OF WEALTH				
		N	D	z-score	95% CI	
PROCEDURE	pooling	76	0.08	0.96	-0.08	0.24
	first decision rule	28	0.14	0.93	-0.38	0.43
	second decision rule	48	-0.15	-1.31	-0.23	0.075
	simultaneous delegation					
	delegation level	170	-0.02	-0.37	-0.14	0.09
	first decision rule	120	-0.14*	-1.77	-0.28	0.01
	second decision rule	50	0.18	1.49	-0.05	0.40
	sequential delegation	60	-0.10	-0.89	-0.33	0.12
	first decision rule	38	0.07	0.45	-0.23	0.36
	second decision rule	22	0.15	0.80	-0.22	0.52

Independent variable: NUMBER OF ROUND		Dependent variable: GINI COEFFICIENT				
		N	D	z-score	95% CI	
PROCEDURE	pooling	76	0.09	0.95	-0.09	0.26
	first decision rule	28	0.17	1.12	-0.13	0.48
	second decision rule	48	-0.11	-0.92	-0.35	0.13
	simultaneous delegation					
	delegation level	170	0.02	0.4	-0.09	0.13
	first decision rule	120	-0.06	-0.79	-0.21	0.09
	second decision rule	50	0.15	1.25	-0.08	0.37
	sequential delegation	60	-0.09	-0.81	-0.31	0.13
	first decision rule	38	-0.03	-0.19	-0.29	0.24
	second decision rule	22	0.29	1.53	-0.08	0.67

A.13 Gini coefficients of the collective results

In Sec. 5.4 I measure the distribution of wealth in the experiment by using the standard deviation and the Gini coefficient of a selected alternative. The comparison of these two measures increased the validity of my analysis. As the Gini coefficients mirrored the results of the standard deviation I include it here.

Table A.10: Gini coefficient

EXPLANATORY NOTE
The table shows the Gini coefficient of the collectively selected alternative for the different procedures. The observations are separated by procedure and for being the first or second decision rule of a session. In addition, I differentiate under SIM for group and delegation level and under SEQ for first and second stage. The table also shows the Gini coefficient which would result if the group had always chosen the core alternative. Obviously this includes only the observations made with tables containing such equilibrium. Please note that the different procedures lead to different core alternatives. The unit of observation are session averages of the collective decisions per procedure.

PROCEDURE	N	mean	SD	min	max
pooling	10	0.44	0.13	0.19	0.64
first decision rule	2	0.34	0.21	0.19	0.48
second decision rule	8	0.46	0.11	0.31	0.64
always core	9	0.51	0.20	0.08	0.84
simultaneous delegation	8	0.38	0.05	0.31	0.43
first decision rule	6	0.37	0.05	0.31	0.43
second decision rule	2	0.40	0.02	0.39	0.42
always core	8	0.21	0.08	0.08	0.31
delegation level	8	0.36	0.04	0.27	0.40
first decision rule	6	0.35	0.05	0.27	0.40
second decision rule	2	0.38	0.02	0.36	0.39
sequential delegation	8	0.51	0.15	0.32	0.78
first decision rule	5	0.54	0.18	0.33	0.78
second decision rule	3	0.44	0.08	0.38	0.52
always core	8	0.30	0.07	0.19	0.37
delegation level	8	0.34	0.09	0.24	0.50
first stage	8	0.37	0.13	0.27	0.66
second stage	8	0.35	0.18	0.11	0.67

A.14 Standard experimental games used to measure social preferences

EXPLANATORY NOTE

The table shows standard experimental games used to investigate social preferences. For every game it provides a short description of set-up, typical findings and interpretation. The table is based on Levitt and List (2007, p. 155, table 1). Another collection can be found in Camerer and Fehr (2004, p. 62). In addition, Ert et al. (2011, p. 257) provided a proper summary of main deviations from rational choice theory in experimental games.

Table A.11: Experimental games to measure social preferences

Game	Summary	Typical finding	Interpretation
Ultimatum game	A two-stage game where two people, a proposer and a responder, bargain over a fixed amount of money. In the first stage, the proposer offers a split of the money, and in the second stage, the responder decides to accept or reject the offer. If accepted, each player receives money according to the offer; if rejected, each player receives nothing.	Proposer: Majority of offers in the range of 25-50% of fixed amount. Few offers below 5%. Responder: Frequently reject offers below 20% of fixed amount.	Proposer: Fairness Responder: Punish unfair offers: negative reciprocity, fairness, inequity aversion
Dictator game	A variant of the ultimatum game: strategic concerns are absent as the proposer simply states what the split will be and the proposer has no veto power, rendering the proposed split as effective.	Above 60% of subjects pass a positive amount of money, the mean transfer is roughly 20% of the endowment.	Altruism; fairness preferences, such as inequity aversion.
Trust game	A sequential prisoner's dilemma game wherein the first mover decides how much money to pass to the second mover. All money passed is increased by a factor, f> 1, and the second mover then decides how much money to return to the first mover. In this light, the second mover is a dictator who has been given his endowment by the first mover.	Proposer: Average transfer of roughly 50% of endowment. Responder: Repayment is increasing in transfer. Average repayment rate is nearly 50% of transfer.	Proposer: Trust; foresee positive reciprocity Responder: Trustworthiness, positive reciprocity
Gift exchange game	Similar to the trust game, but the money passed by the first mover (often labeled the "wage" or "price" offer), is not increased by a factor, rather it represents a pure lump-sum transfer. Also, the first mover requests a desired effort, or quality, level in return for the "wage" or "price" offer. The second mover then chooses an effort or quality level that is costly to provide, but increases the first mover's payoff.	Proposer: "Wage" or "price" offer is typically greater than the minimum allowed. Responder: Effort or quality increases in "wage" or "price" offer.	Proposer: Trust; foresee positive reciprocity Responder: Trustworthiness, positive reciprocity
Public goods game	Generalization of the prisoner's dilemma. N group members decide simultaneously how much to invest in a public good. The payoff function is given by $P_i = e - g_i + \beta \sum_n g_j$ where e represents initial endowment; g are the tokens which subject i places in the group account; β is the marginal payoff of the public good; and $\sum_n g_j$ is the sum of the n individual contributions. By making $0 < \beta < 1 < n$, the dilemma follows.	Contribution to public good is roughly 50% of endowment in one-shot games. Many contributions approach 0% in latter rounds of multi-period games.	Altruism; fairness preferences, conditional reciprocity

A.15 Predicted probabilities for core solutions at different levels of rationality

The figures compare predictions of the sophisticated and the sincere voting model for different levels of random errors. In the sincere model the probability of individual choices follows the standard conditional logit model. In the sophisticated model the probability of the collective outcome is the likelihood by which at least four players chose the same alternative.

Figure A.8: Predicted probabilities under pooling

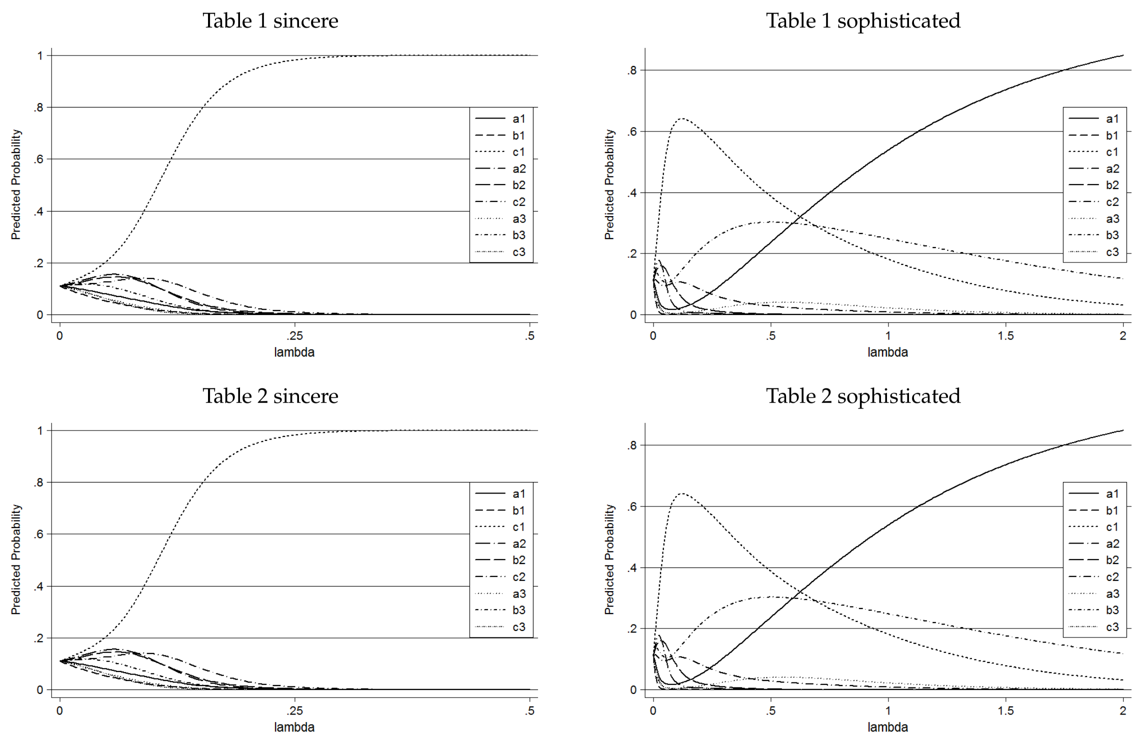


Table 3 sincere

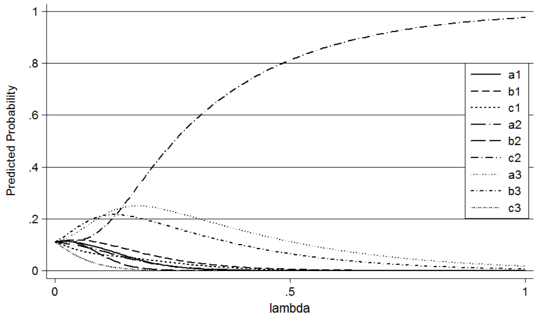


Table 3 sophisticated

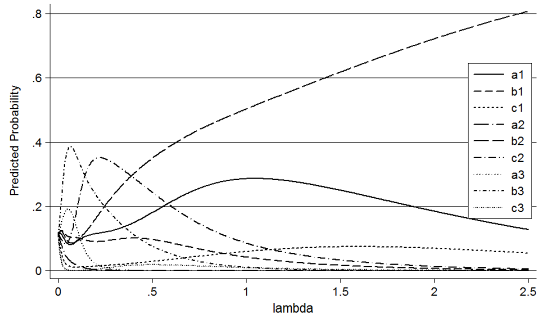


Table 4 sincere

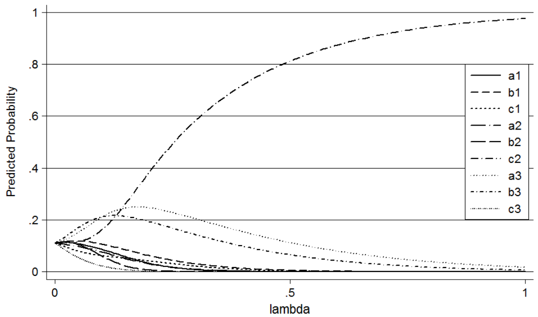


Table 4 sophisticated

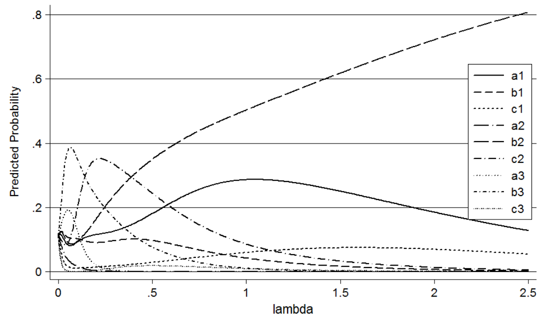


Table 5 sincere

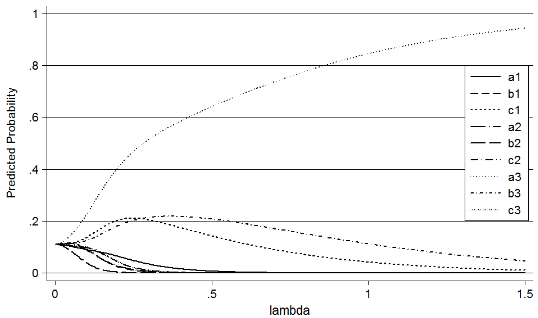


Table 5 sophisticated

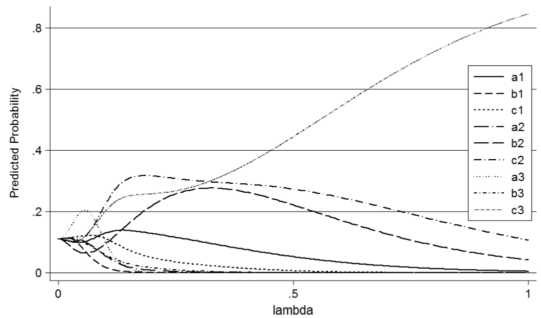


Table 6 sincere

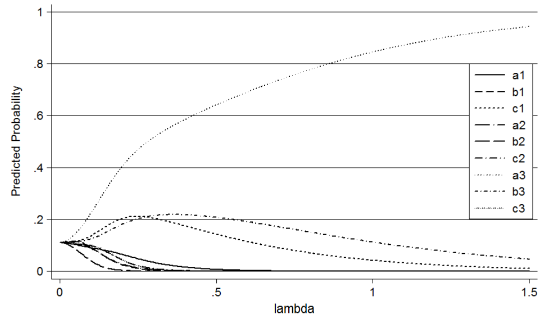


Table 6 sophisticated

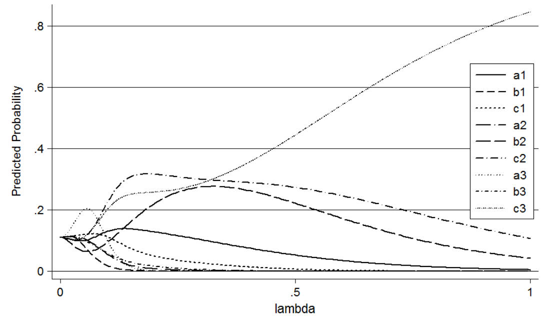


Table 7 sincere

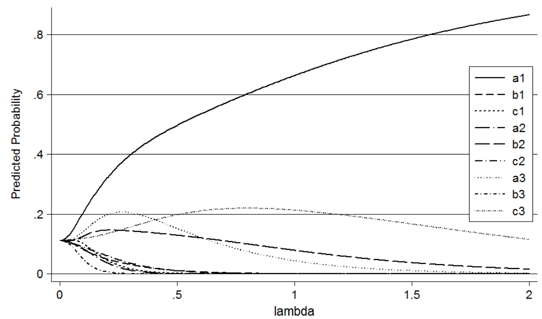


Table 7 sophisticated

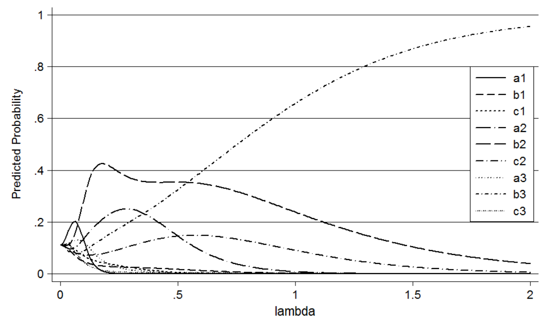


Table 8 sincere

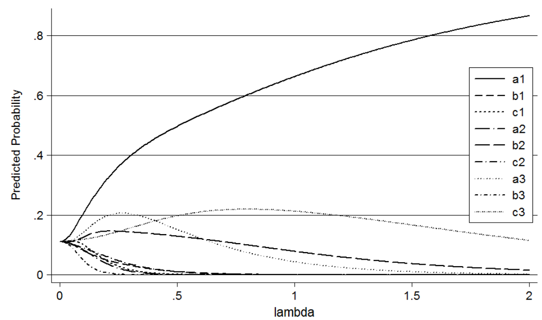


Table 8 sophisticated

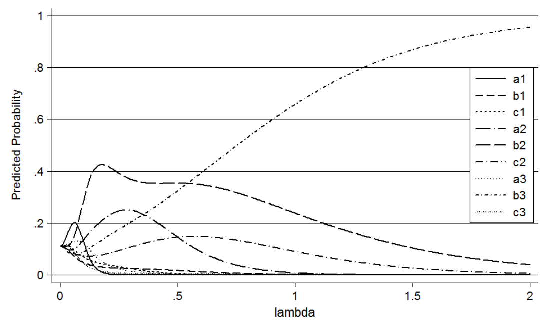


Table 9 sincere

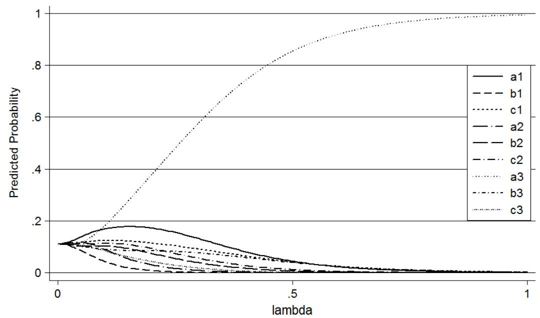


Table 9 sophisticated

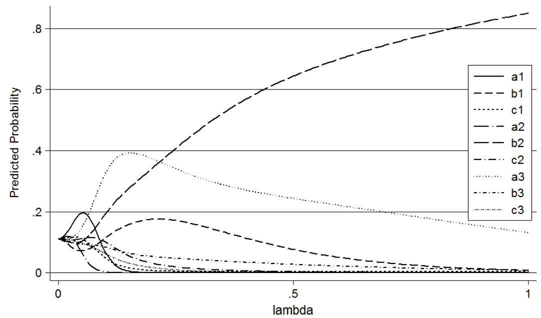


Table 10 sincere

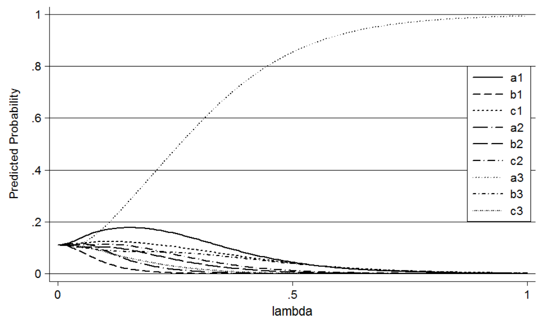


Table 10 sophisticated

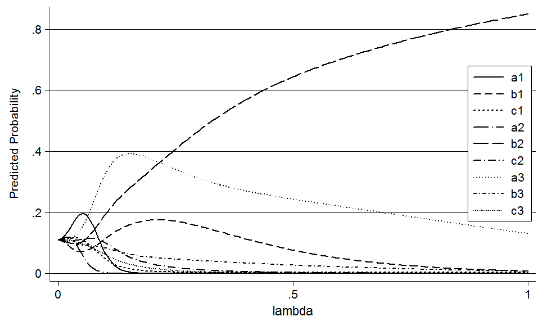


Table 11 sincere

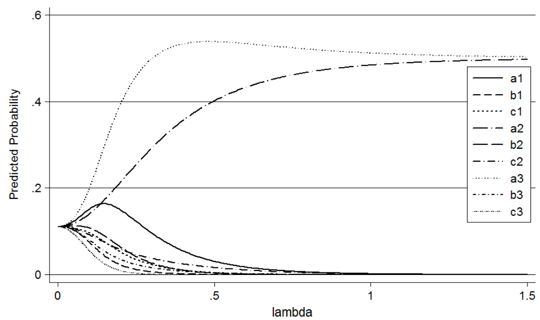


Table 11 sophisticated

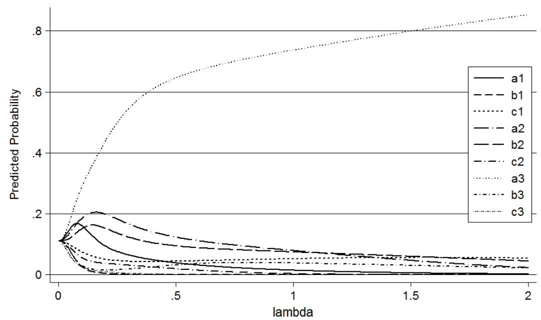


Table 12 sincere

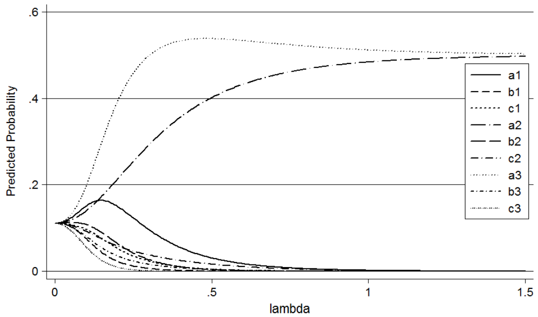


Table 12 sophisticated

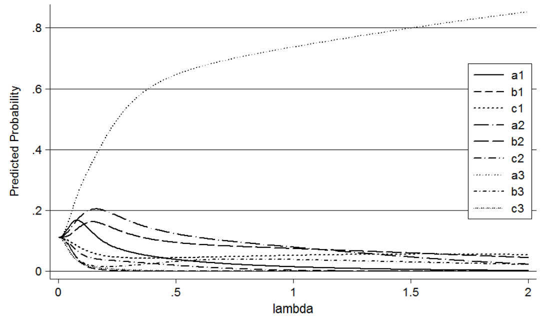


Table 13 sincere

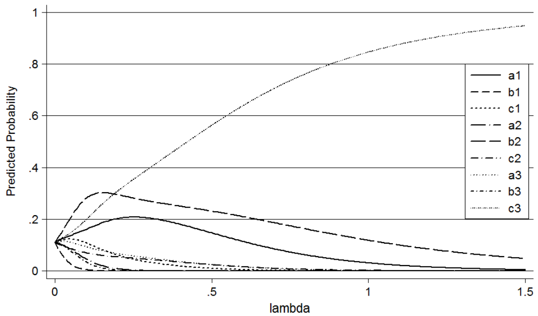


Table 13 sophisticated

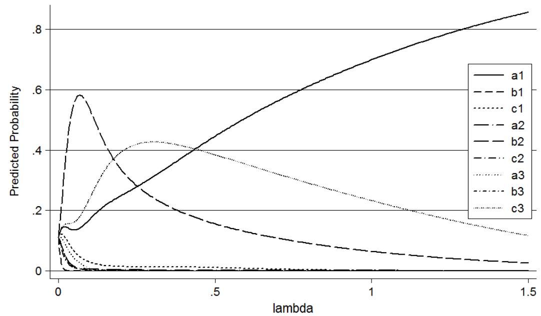


Table 14 sincere

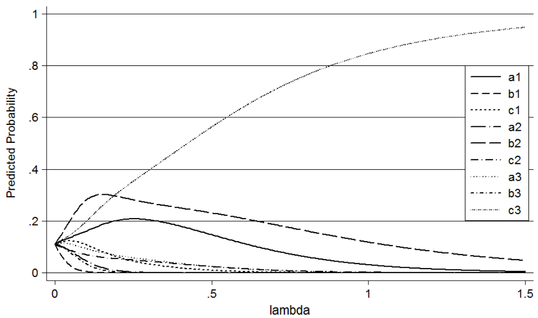


Table 14 sophisticated

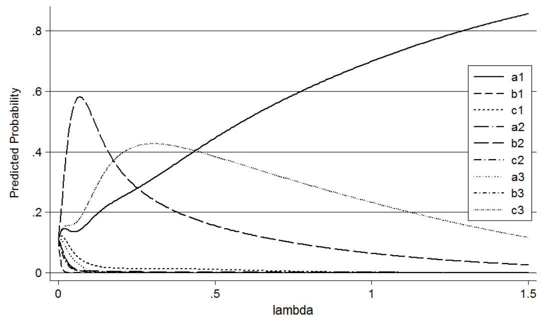


Table 15 sincere

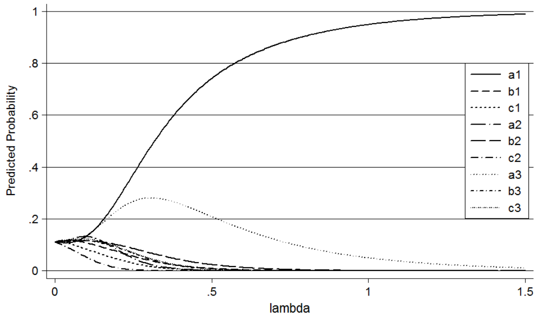


Table 15 sophisticated

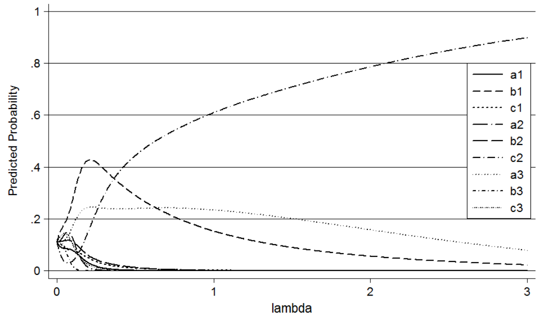


Table 16 sincere

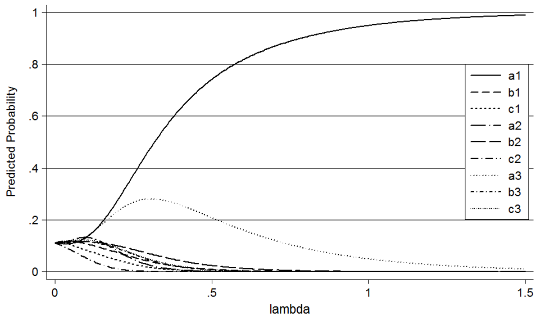


Table 16 sophisticated

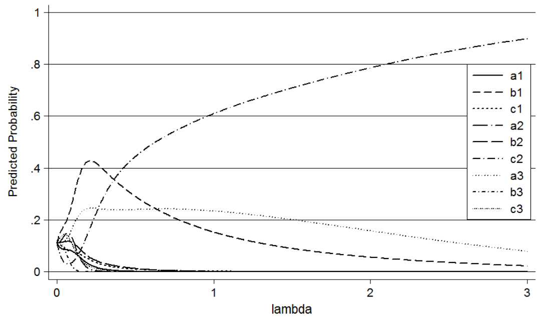


Table 17 sincere

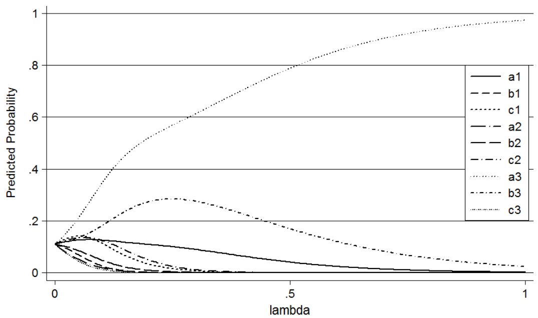


Table 17 sophisticated

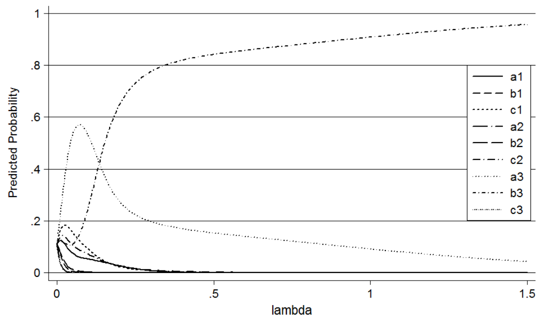


Table 18 sincere

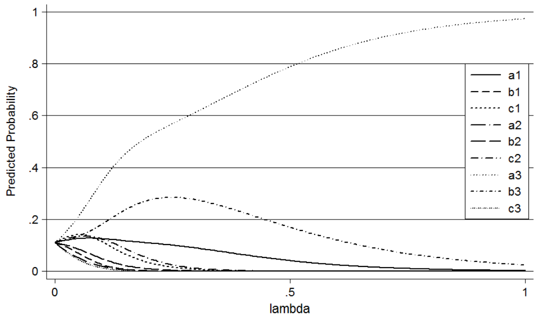


Table 18 sophisticated

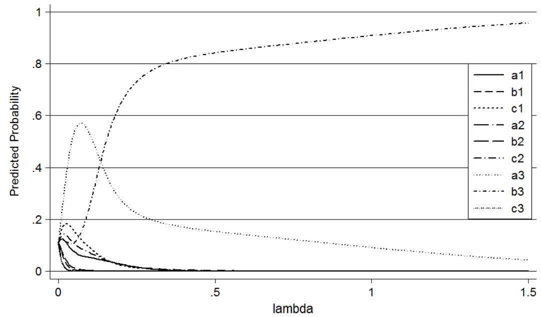


Table 19 sincere

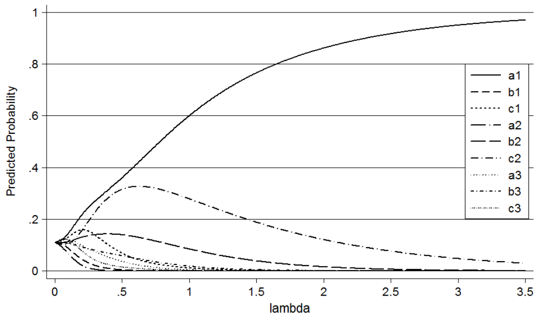


Table 19 sophisticated

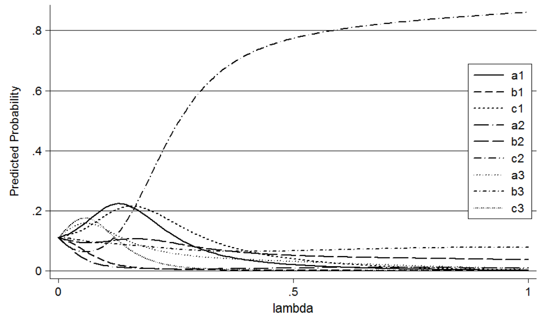


Table 20 sincere

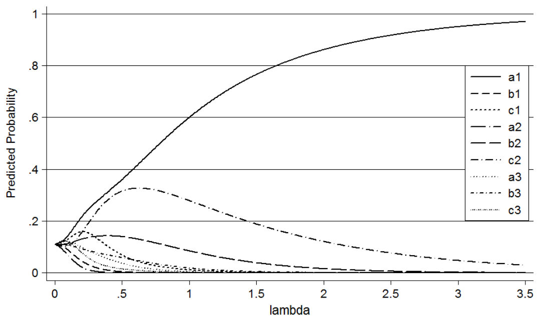


Table 20 sophisticated

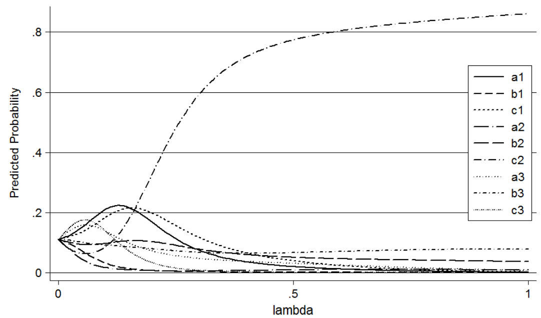


Figure A.9: Predicted probabilities under simultaneous delegation

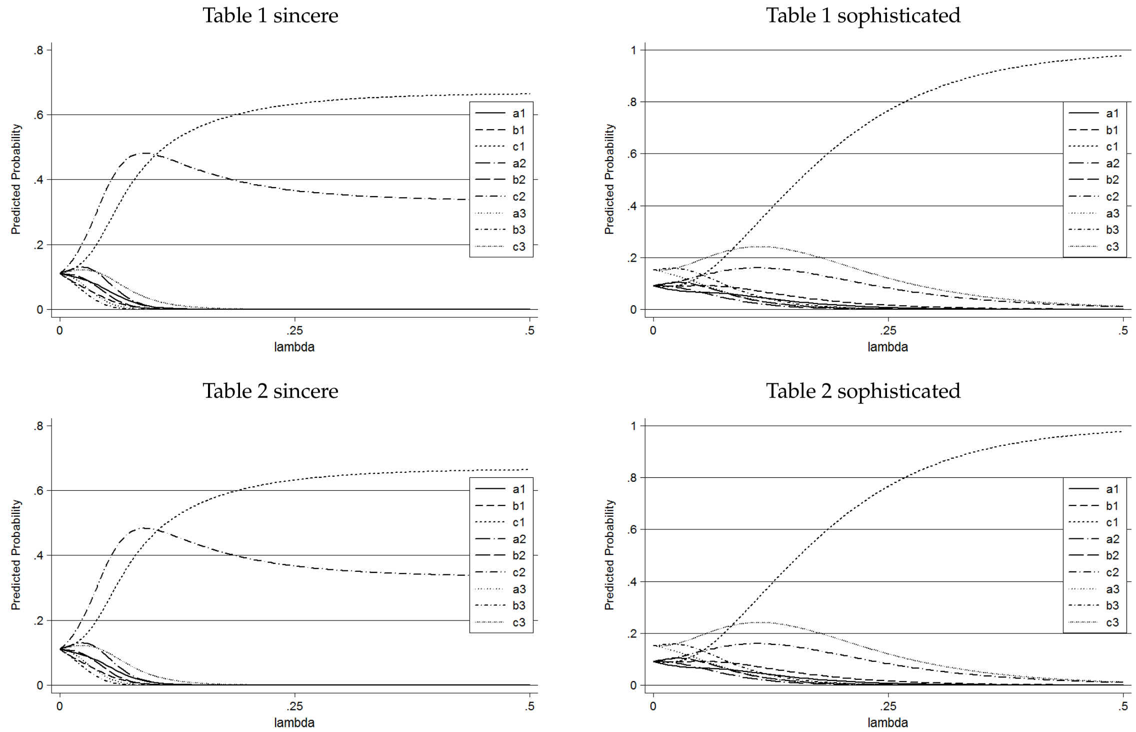


Table 3 sincere

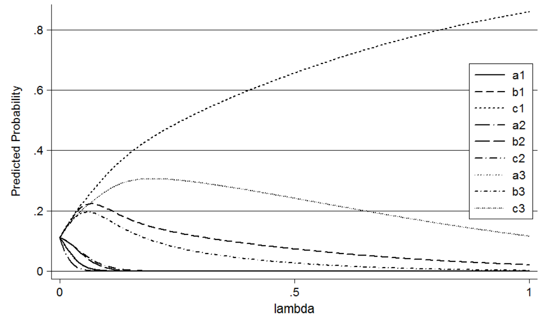


Table 3 sophisticated

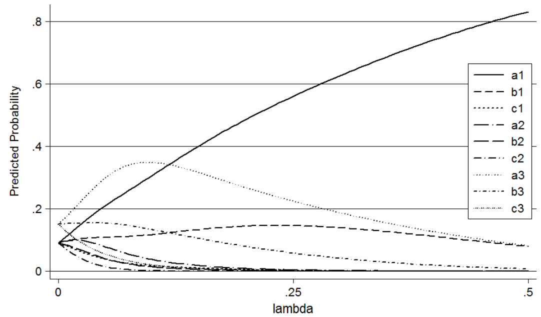


Table 4 sincere

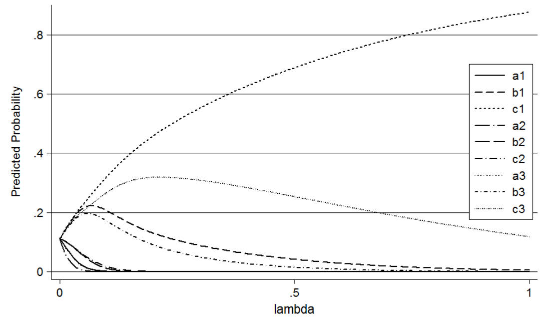


Table 4 sophisticated

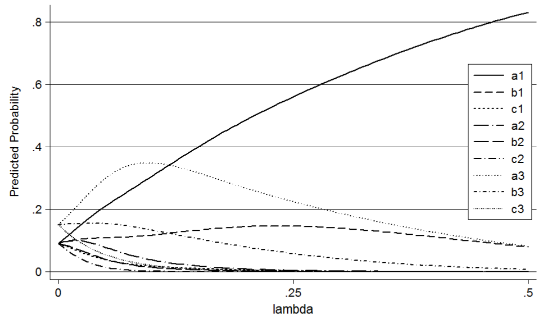


Table 5 sincere

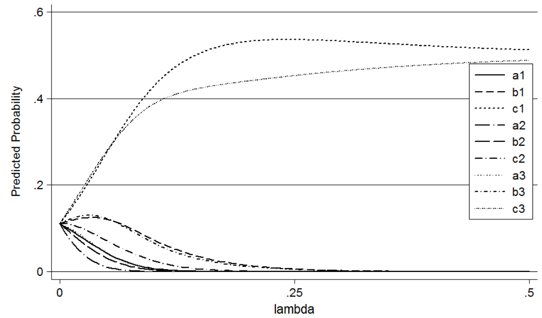


Table 5 sophisticated

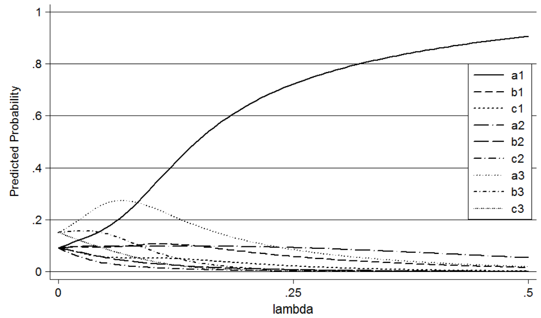


Table 6 sincere

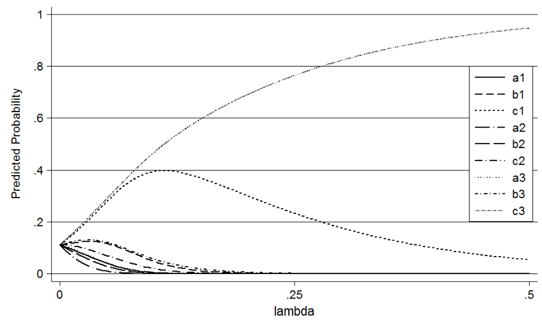


Table 6 sophisticated

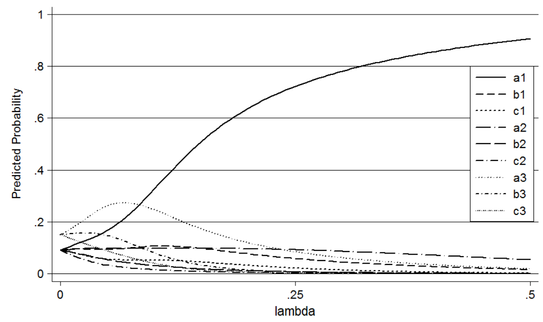


Table 7 sincere

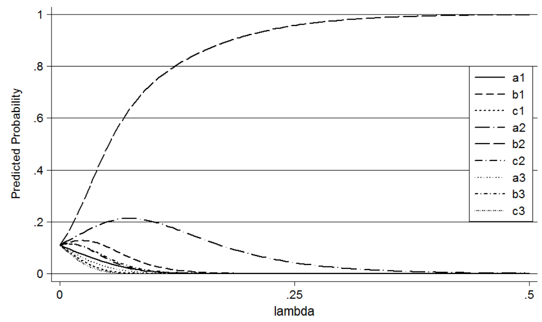


Table 7 sophisticated

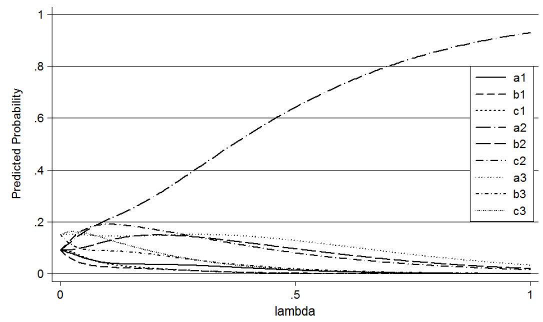


Table 8 sincere

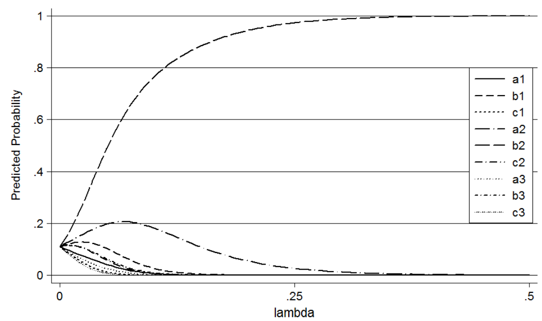
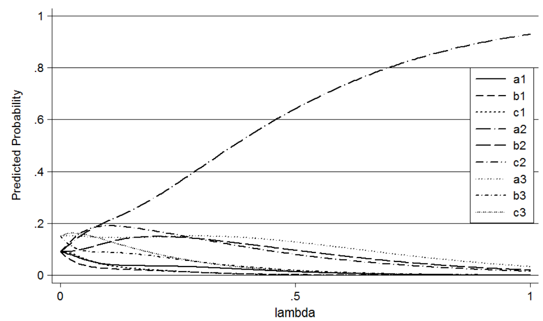


Table 8 sophisticated



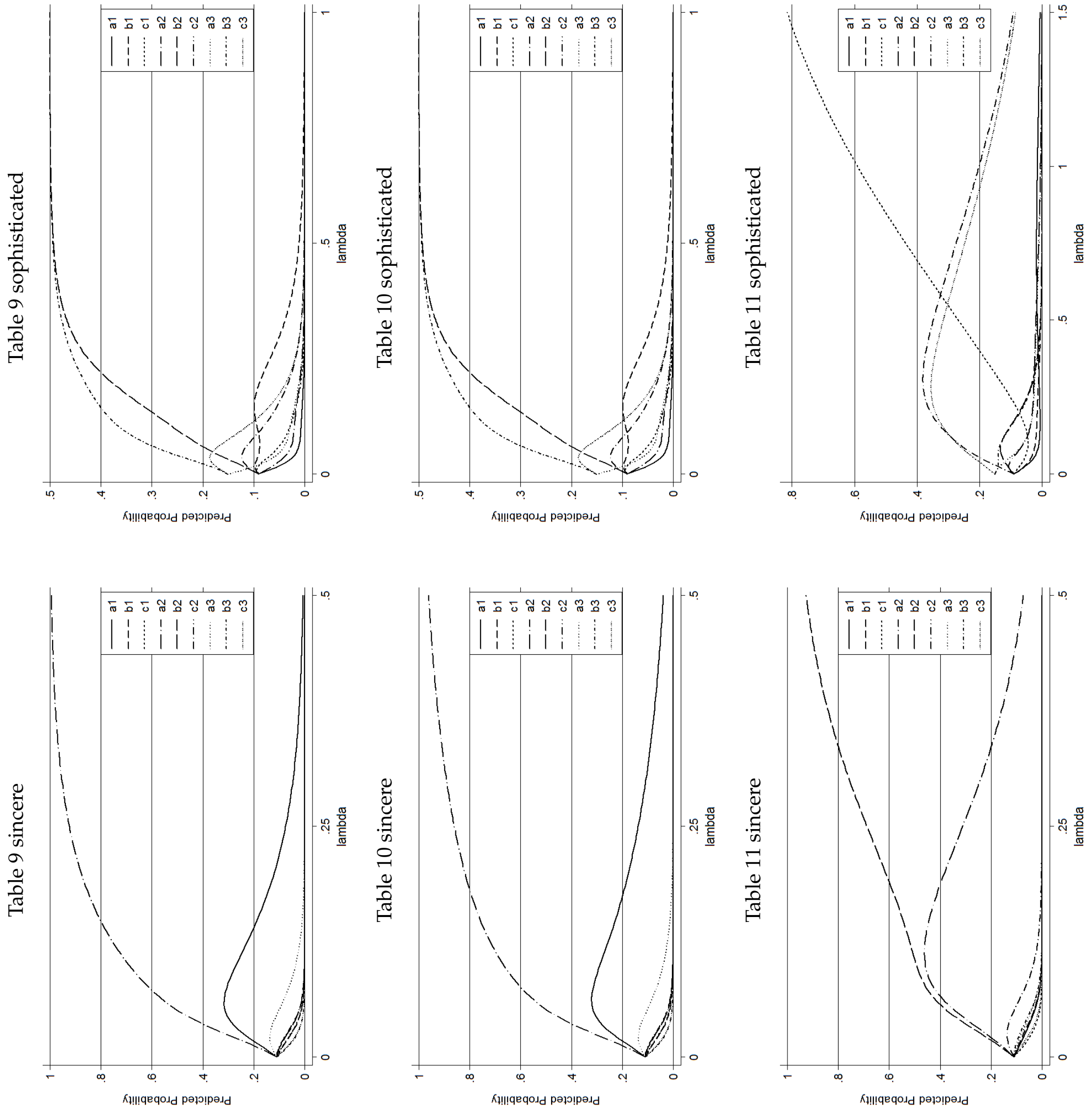


Table 9 sincere
Table 9 sophisticated
Table 10 sincere
Table 10 sophisticated
Table 11 sincere
Table 11 sophisticated
Predicted Probability
lambda
a1
b1
c1
a2
b2
c2
a3
b3
c3

Table 12 sincere

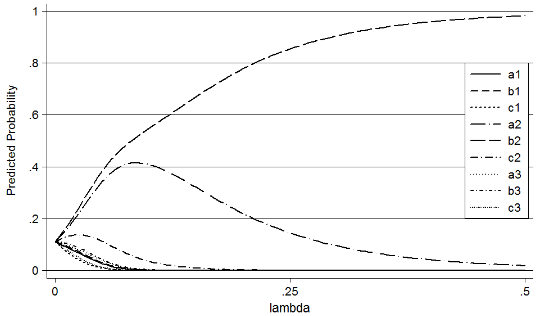


Table 12 sophisticated

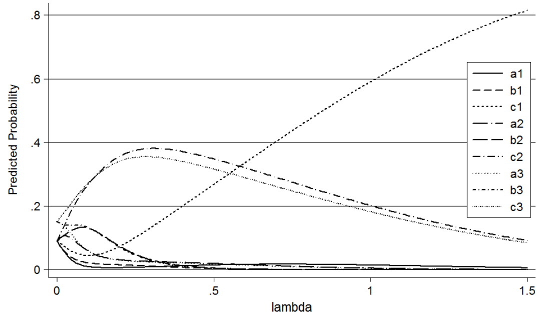


Table 13 sincere

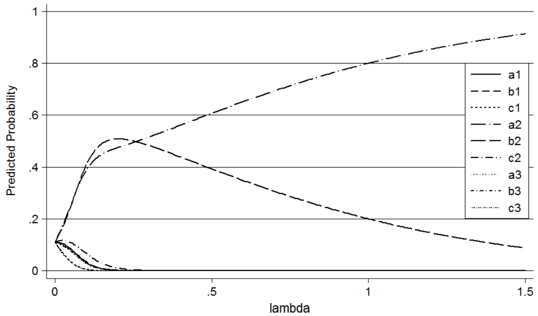


Table 13 sophisticated

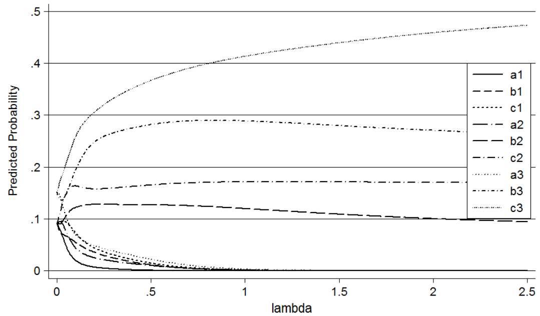


Table 14 sincere

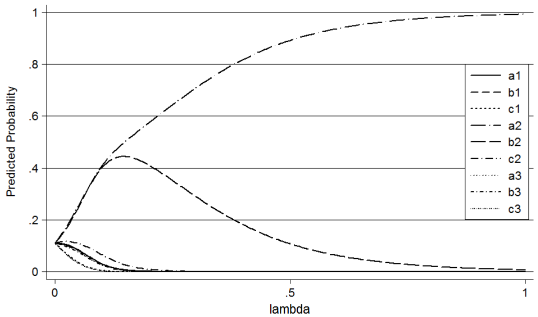


Table 14 sophisticated

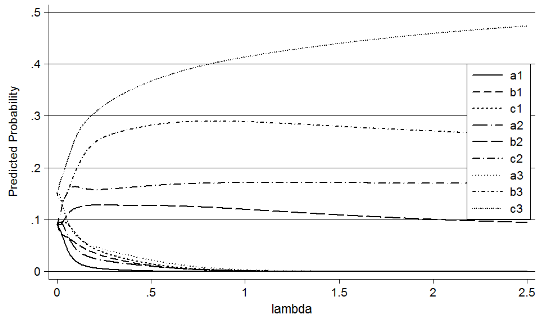


Table 15 sincere

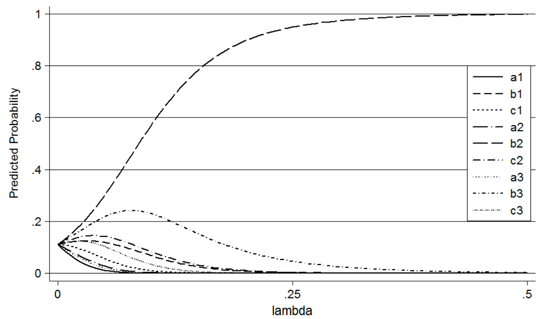


Table 15 sophisticated

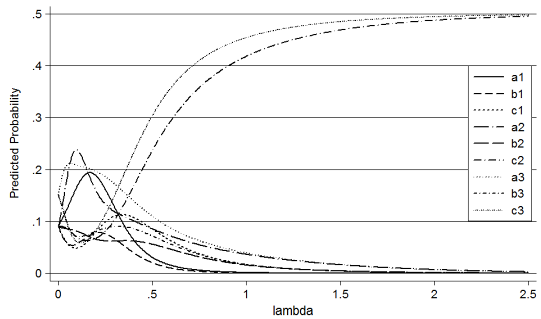


Table 16 sincere

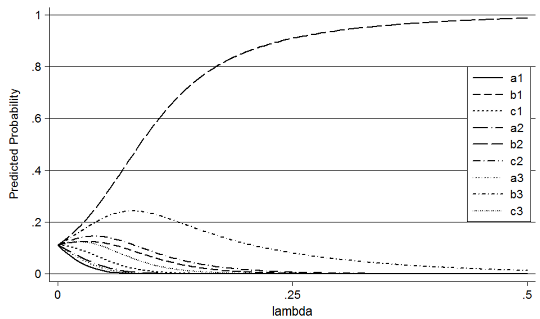


Table 16 sophisticated

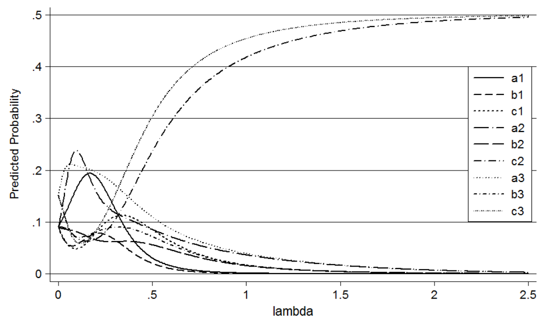


Table 17 sincere

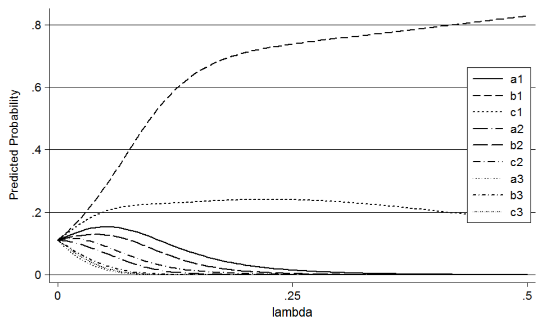


Table 17 sophisticated

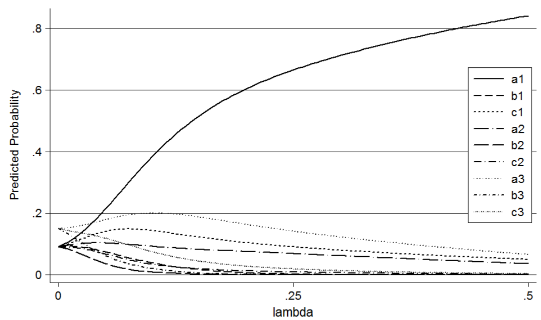


Table 18 sincere

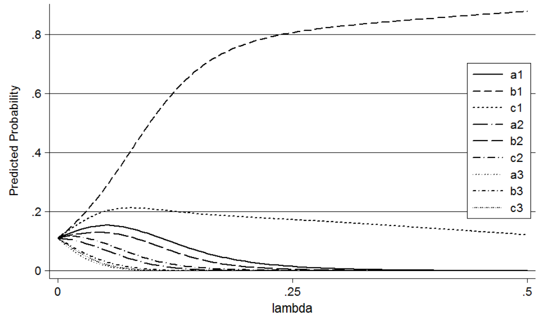


Table 18 sophisticated

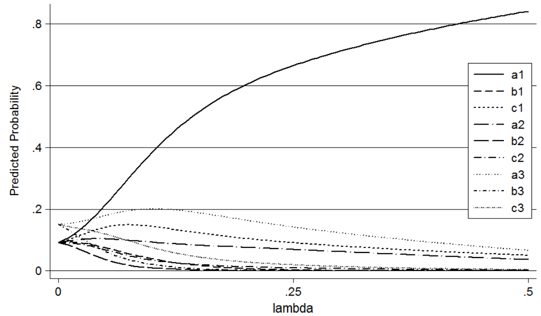


Table 19 sincere

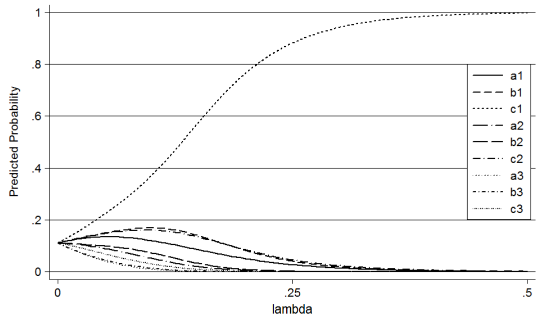


Table 19 sophisticated

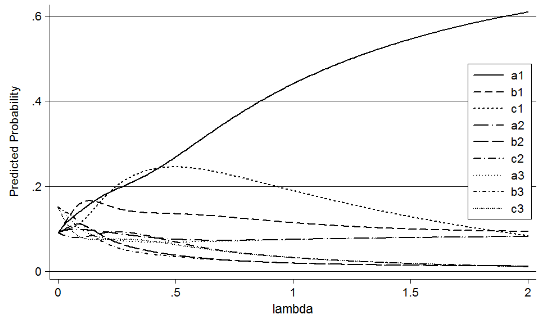


Table 20 sincere

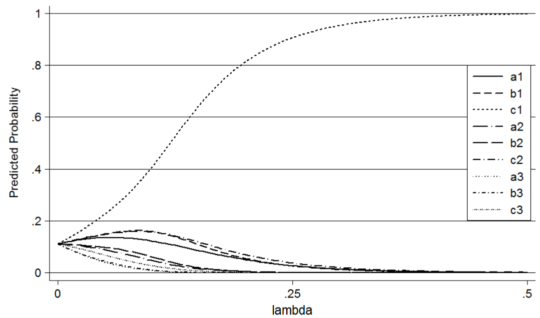


Table 20 sophisticated

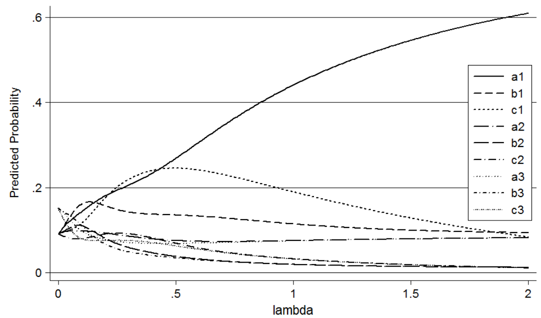


Figure A.10: Predicted probabilities under sequential delegation

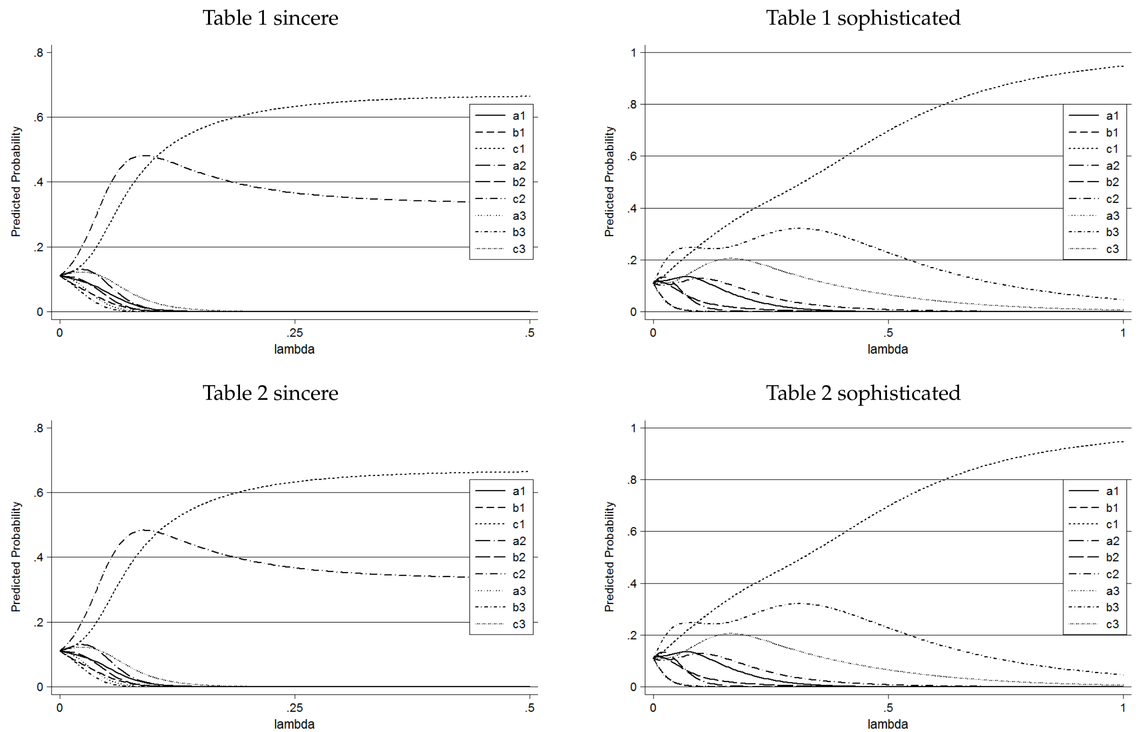


Table 3 sincere

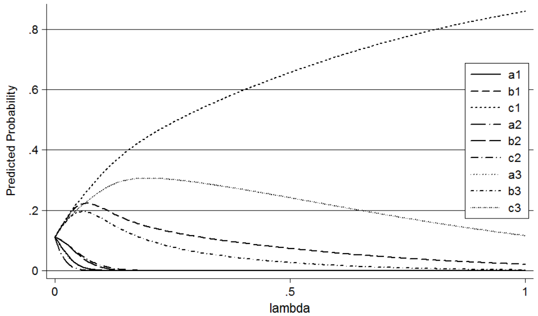


Table 3 sophisticated

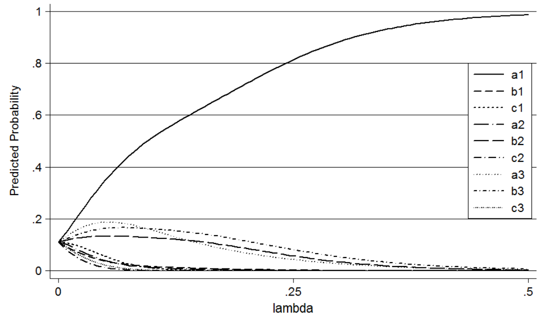


Table 4 sincere

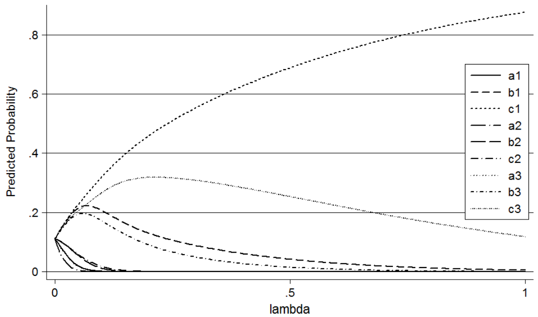


Table 4 sophisticated

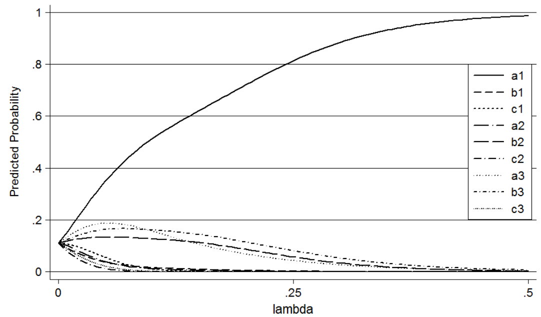


Table 5 sincere

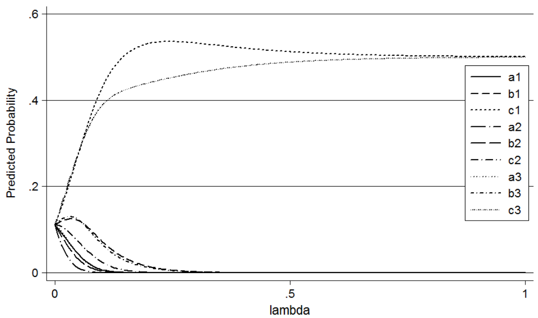
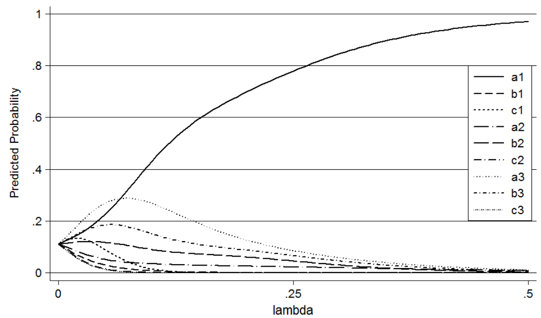


Table 5 sophisticated



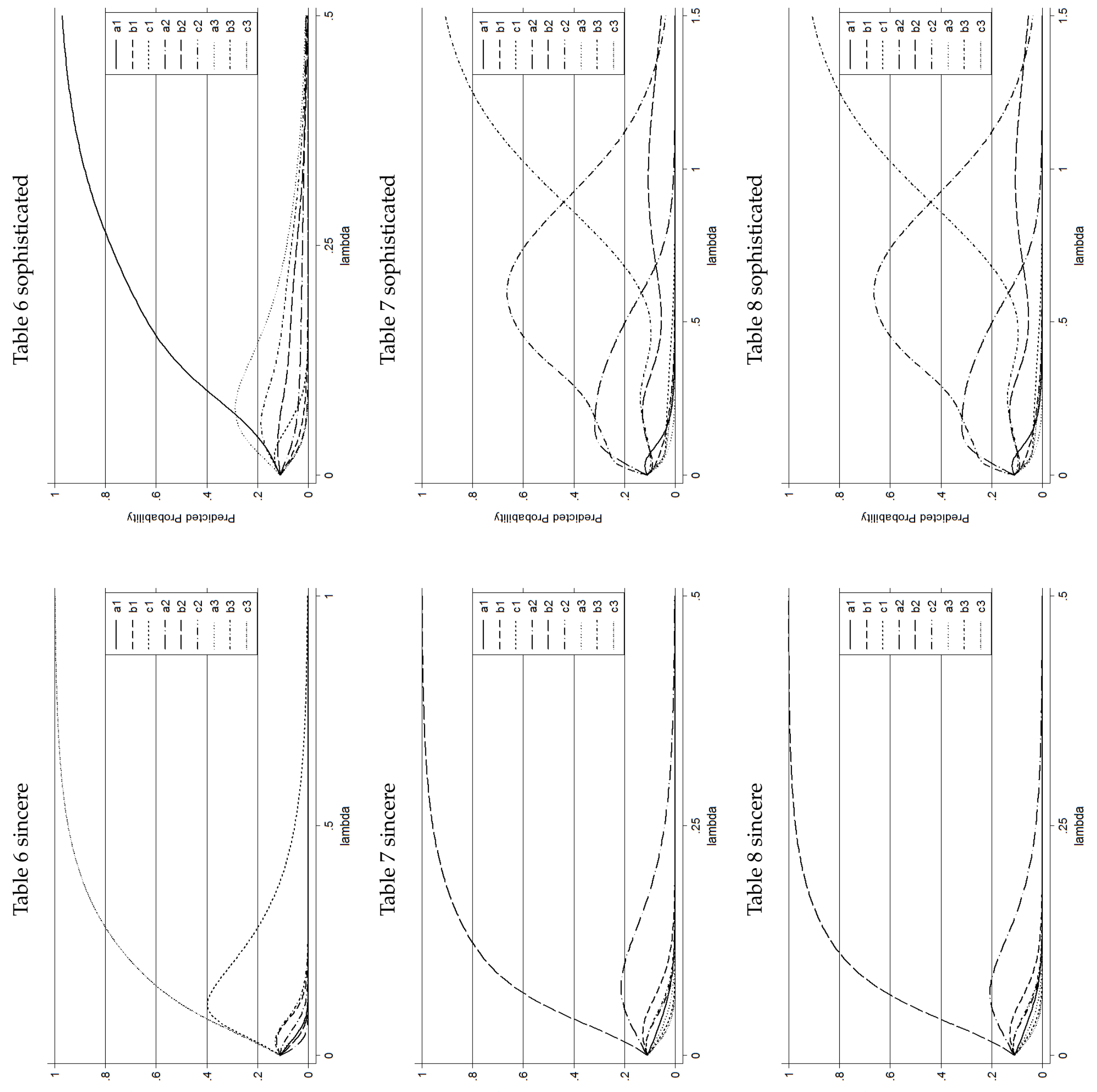


Table 6 sophisticated
Table 7 sophisticated
Table 8 sophisticated
Table 6 sincere
Table 7 sincere
Table 8 sincere
Predicted Probability
lambda
a1
b1
c1
a2
b2
c2
a3
b3
c3

Table 9 sincere

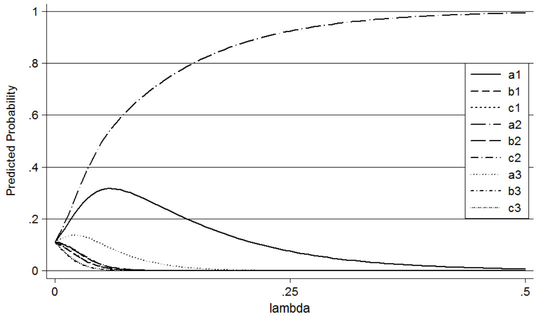


Table 9 sophisticated

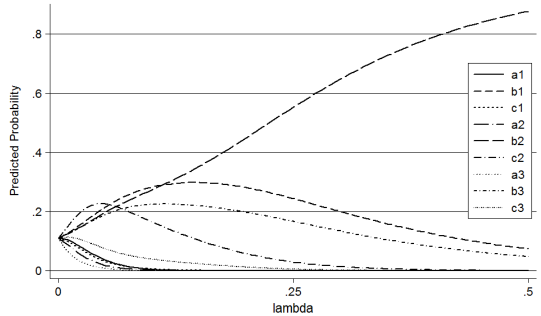


Table 10 sincere

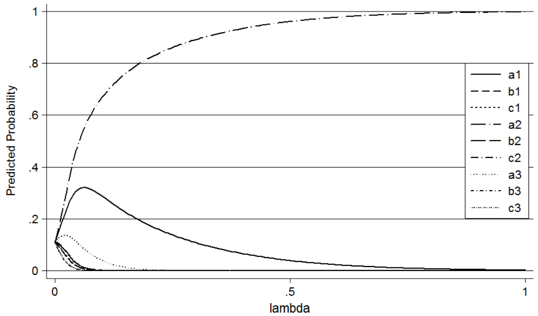


Table 10 sophisticated

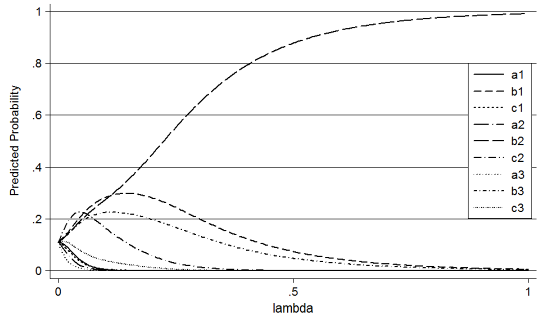


Table 11 sincere

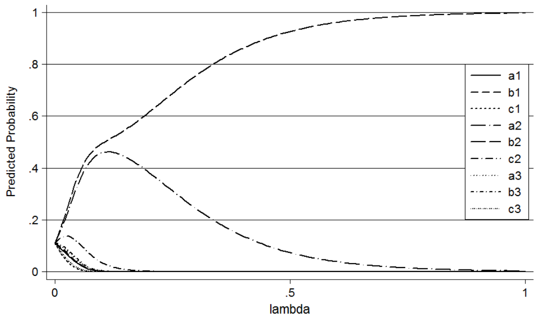


Table 11 sophisticated

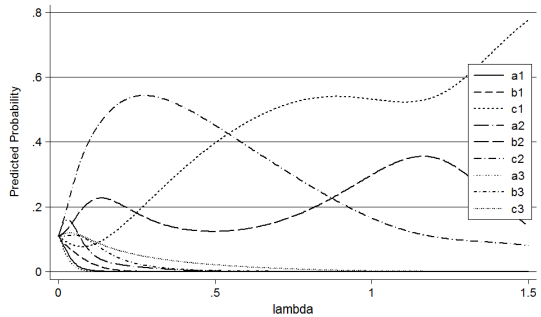


Table 12 sincere

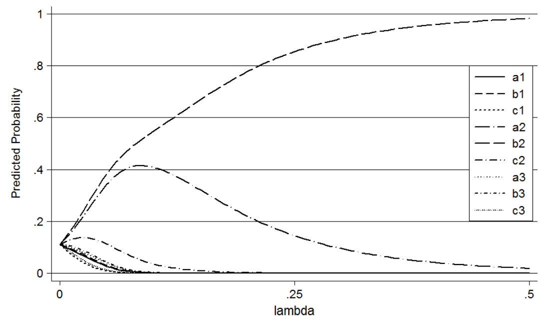


Table 12 sophisticated

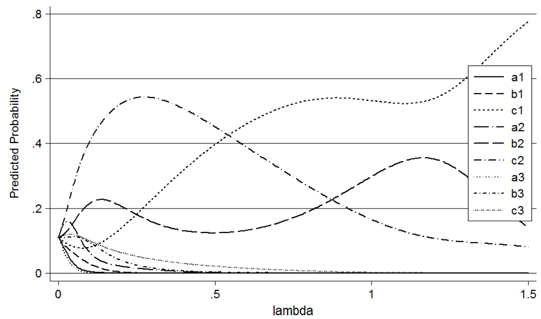


Table 13 sincere

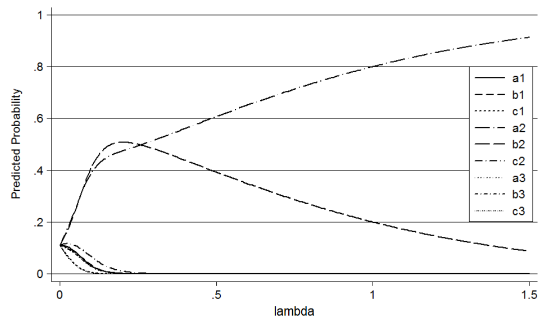


Table 13 sophisticated

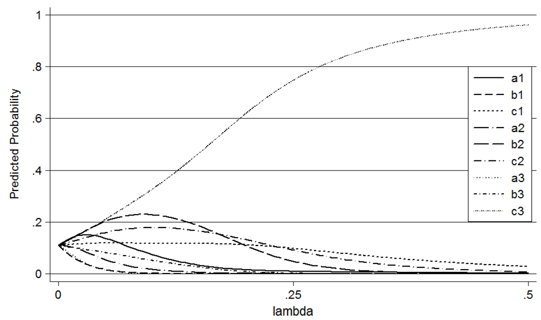


Table 14 sincere

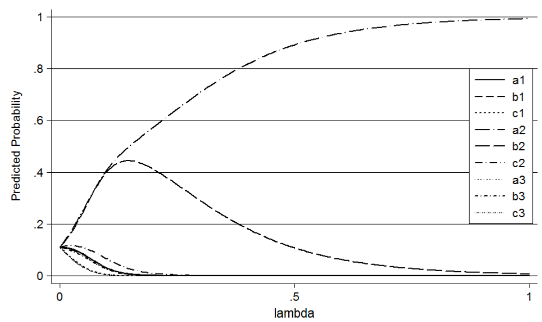


Table 14 sophisticated

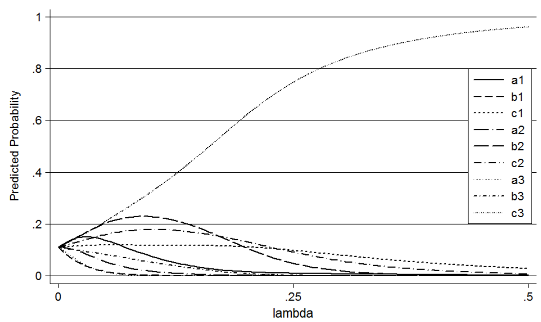


Table 15 sincere

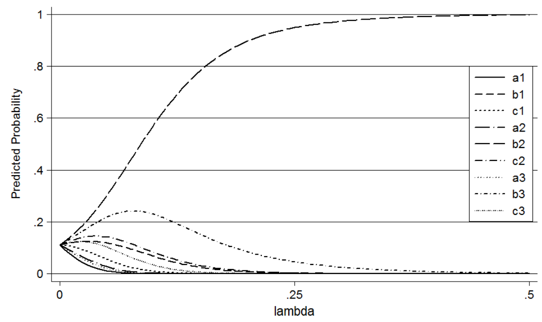


Table 15 sophisticated

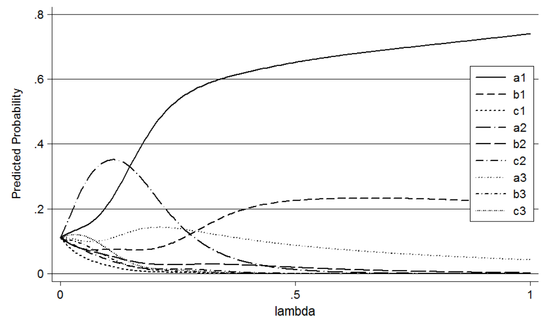


Table 16 sincere

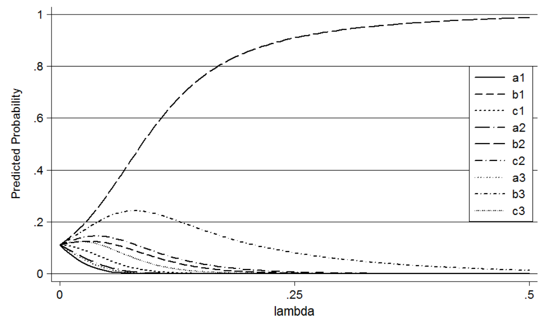


Table 16 sophisticated

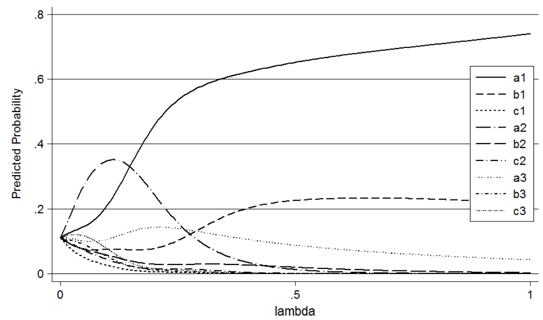


Table 17 sincere

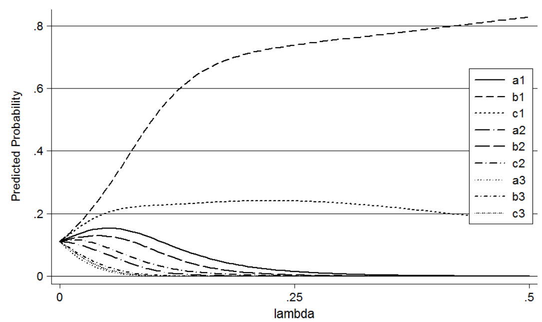


Table 17 sophisticated

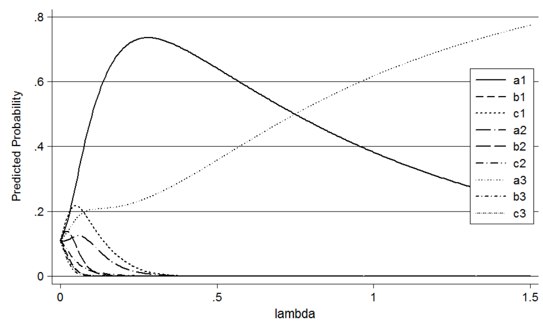


Table 18 sincere

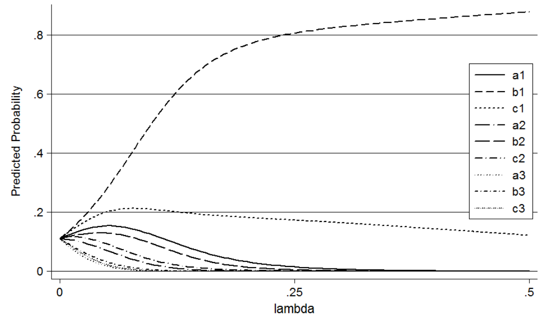


Table 18 sophisticated

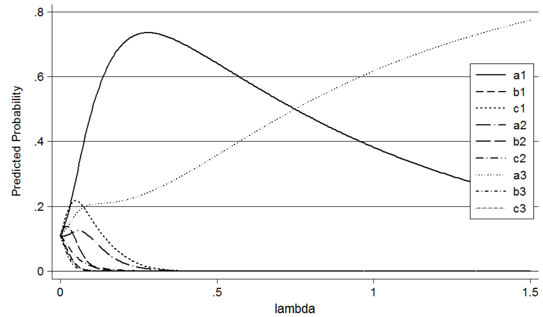


Table 19 sincere

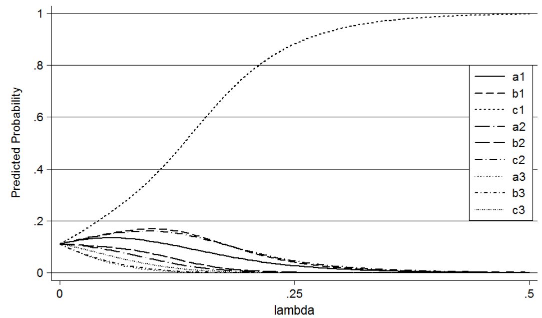


Table 19 sophisticated

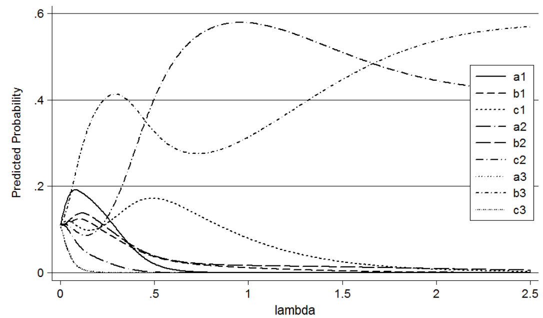


Table 20 sincere

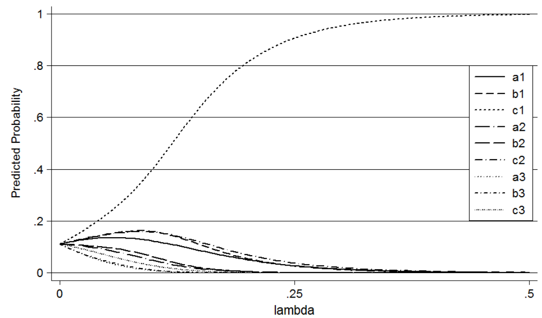
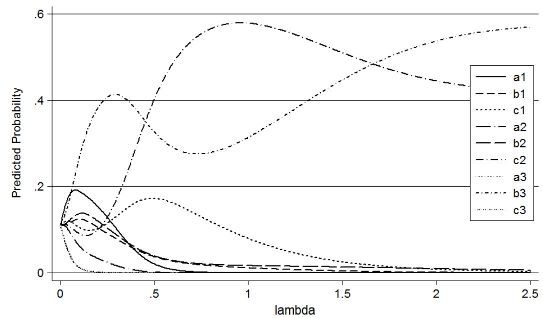


Table 20 sophisticated



A.16 Coding guide for the qualitative content analysis

The following table depicts the coding guide used for the qualitative content analysis in Chap. 7. Its draft followed the approach suggested by Mayring (2002). The guide is only available in German as both the experimental sessions and inquiries were conducted in Heidelberg.

The table states four types of information for every free input field of the post-experiment inquiry. These information are: i) "Kategorie" (German for category), which classifies the type of quotes found in this field; ii) "Definition" which describes the category; iii) "Ankerbeispiel" (German for anchoring example) which plainly exemplifies what kind of quotes are classified into this category; iv) "Kodier-Regel" (German for encoding rule) which points out the norms a quote must fulfill to be sorted into this category.

The information was collected by asking these four questions: i) "Please describe your decision rule with your own words?"; ii) "Which players did you focus on?"; iii) "Which characteristic of an alternative did you consider?"; iv) "What changed when the decision-making procedure split up?". Tabelle A.15 presents the data grouped according to those questions.

Tabelle A.15: Coding guide for the free input fields

Kategorie	Definition	Ankerbeispiel	Kodier-Regel
		Free input field "Please describe your decision rule with your own words?"	
"meine Payoffs"	Der eigene Payoff ist entscheidend.	„Damit meine Auszahlung möglichst hoch wird." „Ich habe in auf die Entscheidung insistiert, die für mich maximale Punktzahl bedeutete."	Nur der eigene Payoff wird explizit genannt.
"Erfolgs-wahrscheinlichkeit meiner Payoffs"	Der eigene Payoff wird mit der Wahlwahrscheinlichkeit gewichtet, mit welcher die anderen Spieler auch diese Alternative wählen.	„Überlegt was die anderen wählen würden und danach meine Wahl ausgerichtet." „Meine beste Auszahlungsmöglichkeit. Dann habe ich berücksichtigt, wie wahrscheinlich es wäre diese auch zu bekommen, wenn man sich die Möglichkeiten der anderen anschaut."	Der strategische Aspekt muss eindeutig benannt werden. Das Ziel ist ein hoher Payoff.
"mein Risiko minimieren"	Risikominimierung der eigenen Auszahlung.	„Möglichst die Spalte gewählt, in welcher für mich keine niedrigen Punkte waren."	Der strategische Aspekt muss eindeutig benannt werden. Das Ziel ist kleine Payoffs zu vermeiden.
"Konsens"	Eine Alternative mit fairer Verteilung, der alle zustimmen können.	„… ein Mittelmaß der Auszahlung zu finden, mit dem eigentlich alle Spieler gut gestellt sind." „Alternativen, die gerecht für alle wären."	Es geht nicht darum das Quorum zu erfüllen (4 von 6). Alle 6 Spieler sollen dem Ergebnis zustimmen können.
"Hohe Summe aller Payoffs"	Maximierung der sozialen Wohlfahrt.	„Die Summe zu maximieren." „Maximaler Nutzen für alle."	Nicht die Verteilung, sondern das Aggregat ist wichtig.

Kategorie	Definition	Ankerbeispiel	Kodier-Regel
		Free input field “Which players did you focus on?”	
“auf mich”	Der eigene Payoff ist entscheidend.	„Ich habe versucht, für mich selbst zuerst zu maximieren.“ „Ich habe versucht viele Punkte zu sammeln und deswegen habe ich auf mich geachtet.“	Nur der eigene Payoff wird explizit genannt.
“auf mein Team”	Die Payoffs des eigenen Teams werden berücksichtigt.	„Hauptsächlich auf die Leute in meiner Gruppe.“ „Ich habe versucht mit meiner Entscheidungen eine Wahl zu treffen, bei der mein Team gut punktet.“	Nur das eigene Team wird genannt.
“auf das andere Team”	Die Payoffs des anderen Teams werden berücksichtigt.	„Ich habe geguckt, was die zweite Gruppe wahrscheinlich wählt.“ „Ich habe darauf geachtet welche Zeile das andere Team ausgehend von ihren Gewinnchancen wohl wählen wird.“	Nur das andere Team wird genannt.
“alle sechs Spieler”	Die Payoffs aller Spieler werden berücksichtigt.	„Die Auswahl bezog sich in erster Linie auf meine Gruppe. Danach sah ich mir möglichen Entscheidungen der anderen Gruppe an.“	Beide Teams und / oder alle sechs Spieler werden genannt.
“genug Spieler in beiden Teams, welche für diese Alternative stimmen”	Einzelne Spieler in beiden Teams werden genannt.	„Ich habe jeweils die Strategien der anderen eingeschätzt.“ „Welche Präferenzen andere Spieler haben werden.“	Das Ziel ist das Erreichen des Quorums (2 mal 2 von 3 Spielern).
		Free input field “Which characteristic of an alternative did you consider?”	
“meine Payoffs”	Der eigene Payoff ist entscheidend.	„Ich habe auf die Höhe meines Payoffs geachtet.“ „Das eigene Ergebnis war entscheidend und nicht das der anderer.“	Nur der eigene Payoff wird explizit genannt.
“mindestens das Zweitbeste meiner Möglichkeiten”	Verzicht auf den höchsten Payoff um einen hohen Payoff für sich zu sichern.	„… es sei denn es gab eine Alternative, die für mich nur eine gering schlechtere Auszahlung bedeutete.“ „Damit die Wahrscheinlichkeit einen hohen Payoff zu erhalten am größten war, habe ich nicht immer den höchsten Payoff gewählt.“	Der strategische Verzicht muss eindeutig benannt werden.
“mein Risiko minimieren”	Risikominimierung der eigenen Auszahlung.	„… nicht zu hohes Verlustrisiko.“ „Ich habe geschaut, was für mich dabei rausspringt, aber auch, dass ich keinen großen Verlust machen konnte.“	Der strategische Aspekt muss eindeutig benannt werden. Das Ziel ist einen geringen Payoff zu vermeiden.
“Erfolgs-wahrscheinlichkeit meiner Payoffs”	Der eigene Payoff wird mit der Wahlwahrscheinlichkeit gewichtet, mit welcher die anderen Spieler auch diese Alternative wählen.	„Zwecks Entscheidungsfindung geschaut, wie die Payoffs für alle wären und welche Alternative so die am wahrscheinlichste Wahl wird.“	Der strategische Aspekt muss eindeutig benannt werden. Das Ziel ist das Erreichen eines hohen Payoffs.

Kategorie	Definition	Ankerbeispiel	Kodier-Regel
“mindestens vier hohe Payoffs”	Erreichen des Quorums.	„Ich habe eine Möglichkeit gewählt, bei der mindestens 4 Personen hohe Auszahlungen hatten.“ „Ich habe geschaut, welche Alternative für mich die Beste gewesen wäre, und dann, ob sie denn auch für 3 andere noch in Frage käme“.	Die Verteilung und Summe aller Payoffs ist nicht entscheidend. Es wird auf das Quorum fokussiert.
“Differenz”	Differenz zwischen dem höchsten und dem kleinstem Payoff.	„Einer ganz hohen und geringen Auszahlung bin ich möglichst aus dem Weg gegangen.“	Höchster und kleinster Payoff sind entscheidend.
“Verteilung”	Möglichst gleiche Verteilung.	„… aber möglichst niemand extrem wenige Punkte bekommt.“ „…versucht geringe Auszahlungen für Spieler zu vermeiden.“	Bei der Verteilung werden alle sechs Spieler berücksichtigt.
Free input field “What changed when the decision-making procedure split up?”			
“es änderte sich nichts”	Keine Veränderung.	„kein Unterschied“ „Gleiches Prinzip“ „Viel mehr auf meine eigenen Payoffs, vor allem ein möglichst hohes Minimum.“	Nur der eigene Payoff wird explizit genannt.
“mehr Egoismus”	Der eigene Payoff wird wichtiger.	„Da habe ich hauptsächlich nur auf meine Punkte geachtet, da die Entscheidung eingeschränkt wurde.“	Veränderung muss explizit genannt werden. Betonung des eigenen Payoffs.
“mehr Kompromiss”	Die Zustimmung aller Spieler wird wichtiger.	„Es ging weniger um den eigenen Vorteil.“ „Ich habe eher Kompromisse gemacht.“	Veränderung muss explizit genannt werden. Konsens muss betont werden.
“mehr mein Team im Vordergrund”	Die Teammitglieder sind jetzt wichtiger.	„Mehr auf die Auszahlung meiner Teammitglieder.“ „… habe ich mehr auf meine Teammitglieder bzw. deren Punktzahlen geachtet.“	Veränderung muss explizit genannt werden. Abgrenzung des eigenen Teams nötig.

A.17 References of experiments conducted online and in the laboratory

EXPLANATORY NOTE

In December 2012 Israel Waichman (Department of Economics, Heidelberg University) conducted a survey via the Economic Science Association (ESA) Google Group discussion list "ESA Experimental Methods Discussion" [461]. He asked all subscribers for references to studies comparing between experiments conducted in the laboratory and via the internet. The listing below resembles his posted results (Google Group entry from Wednesday 2nd January 2013 13:15).

Amir, O., D. G. Rand, and Y. K. Gal. 2012. "Economic games on the internet: The effect of $1 stakes". *PLOS ONE* 7.

Anderhub, V., R. Müller, and C. Schmidt. 2001. "Design and evaluation of an economic experiment via the internet". *Journal of Economic Behavior and Organization* 46:227-247.

Bellemare, C. and S. Kröger. 2007. "On Representative Social Capital". *European Economic Review*, 51:183-202.

Birnbaum, M. H. 2000. *Psychological Experiments on the Internet.* Academic Press.

Bolton, G., B. Greiner, and A. Ockenfels. 2012. "Engineering Trust: Reciprocity in the Production of Reputation Information". *Management Science*.

Bosch-Domènech, A., J. García-Montalvo, R. Nagel, and A. Satorra. 2002. "One, Two, (Three), Infinity…: Newspaper and Lab Beauty-Contest Experiments". *American Economic Review* 92:1687-1701.

Bosch-Domènech, A., and R. Nagel. 1997 "Cómo se el da la bolsa," *Expansión*, June 4:40.

Bosch-Domènech, A., and R. Nagel. 1997. "El juego de adivinar el numero X: una explicación y la proclamación del vencedor," *Expansión*, June 16:42-43.

Bosch-Domènech, A., and R. Nagel. 1997. "Guess the Number: Comparing the FT's and Expansion's Results," *Financial Times*, Section Mastering Finance 8, June 30:14.

Buchanan, T. and J. L. Smith. 1999. "Using the internet for psychological research: personality testing on the World Wide Web". *British Journal of Psychology* 90:125-144.

Cassar, A. 2007. "Coordination and cooperation in local, random and small world networks: Experimental evidence". *Games and Economic Behavior* 8:209-230.

Cabrales, A., and R. Nagel. 2002. "Ein Gewinnspiel zur Entscheidung unter Unsicherheit". *Spektrum der Wissenschaft* (German issue of Scientific American) Nov.

Cabrales, A., and R. Nagel. 2003. "Entscheidung unter Unsicherheit – die Ergebnisse". *Spektrum der Wissenschaft* (German issue of Scientific American) May.

Charness, G., E. Haruvy, and D. Sonsino. 2007. "Social distance and reciprocity: An internet experiment". *Journal of Economic Behavior & Organization* 63:88-103.

Chesney, T., S.-H. Chuah, and R. Hoffmann. 2009. "Virtual world experimentation: An exploratory study". *Journal of Economic Behavior & Organization* 72:618-635.

Cleave, Blair L., N. Nikiforakis, and R. Slonim. 2012. "Is there Selection Bias in Laboratory Experiments? The Case of Social and Risk Preferences". *Experimental Economics* (forthcoming).

[461] The discussion can be viewed after prior registration under `https://groups.google.com/d/msg/esa-discuss/-/iTvrjXqCod4J`.

Coffman, L. C. 2011. "Intermediation reduces punishment (and reward)". *American Economic Journal: Microeconomics* 3:77-106.

Costa, Gomes M., V. Crawford, and R. Nagel. 2007. "Experiment teilzunehmen". *Spektrum der Wissenschaft* (German issue of Scientific American) July:96-97.

Costa, Gomes M., V. Crawford, and R. Nagel. 2008. "Ergebnisse zum Preisausschreiben". *Spektrum der Wissenschaft* (German issue of Scientific American) February:76-80.

Dürsch, P., J. Oechssler, and B. C. Schipper. 2009. "Incentives for subjects in internet experiments". *Economics Letters* 105:120-122.

Dutcher, E. Glenn. 2012. "The effects of telecommuting on productivity: An experimental examination. the role of dull and creative tasks". *Journal of Economic Behavior & Organization* 84:355-363.

Dutcher, E. Glenn and Krista Jabs Saral. 2012. "Does team telecommuting affect productivity? An experiment". *mimeo*. https://sites.google.com/site/glenndutcher/Home/research?pli=1.

Eckel, C. C., and R. K. Wilson. 2006. "Internet Cautions: Experimental Games with Internet Partners". *Experimental Economics* 9:53-66.

Egas, M. and A. Riedl. 2008. "The economics of altruistic punishment and the maintenance of cooperation". *Proceedings of the Royal Society B* 275:871-878.

Fiedler, M. and E. Haruvy. 2009. "The lab versus the virtual lab and virtual field - an experimental investigation of trust games with communication". *Journal of Economic Behavior and Organization* 72(2):716-724.

Füllbrunn, S., K. Richwien, and A. Sadrieh. 2011. "Trust and trustworthiness in anonymous virtual worlds". *Journal of Media Economics* 24:48-63.

Gosling, Samuel D., Simine Vazire, Sanjay Srivastava, and Oliver P. John. 2004. "Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions about Internet questionnaires". *American Psychologist* 59: 93-104.

Jerit, J., J. Barabas, and S. Clifford. 2012. "Comparing Contemporaneous Laboratory and Field Experiments on Media Effects". *mimeo. Department of Political Science at the Florida State University*. http://scottaclifford.com/wp-content/uploads/2012/12/JeritBarabasClifford_Comparing_Experiments.pdf.

Hergueux, J. and N. Jacquemet. 2012. "Social preferences in the online laboratory: A randomized experiment". *CES working papers 2012* 70.

Ho, B., J. Taber, G. Poe, and A. Bento. 2012. "Moral Licensing and Moral Cleansing in Contingent Valuation and Laboratory Experiments of Willingness to Pay to Reduce Negative Externalities". *mimeo. Vassar College, Economics Department*. http://idei.fr/doc/conf/bee2012/Culpability%20CV%20experimental%206-8-12.pdf.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser. 2011. "The online laboratory: conducting experiments in a real labor market". *Experimental Economics* 14:399-425.

Lyons, B., G. Menzies, and D. Zizzo. 2012. "Conflicting evidence and decisions by agency professionals: an experimental test in the context of merger regulation". *Theory and Decision* 73:465-499.

Nagel, R., and R. Selten. 1997. "1000 DM zu gewinnen". *Spektrum der Wissenschaft* (German issue of Scientific American) Nov.

Nagel, R., and R. Selten. 1998 "Das Zahlenwahlspiel - Hintergruende und Ergebnisse" *Spektrum der Wissenschaft* (German issue of Scientific American) Febr.:16-22.

Paolacci, G., J. Chandler, and P. G. Ipeirotis. 2010. "Running experiments on Amazon mechanical turk". *Judgment and Decision Making* 5:411-419.

Potters, J. and S. Suetens. 2009. "Cooperation in experimental games of strategic complements and substitutes". *Review of Economic Studies* 76:1125-1147.

Rand, D. G. 2012. "The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments". *Journal of Theoretical Biology* 299: 172-179

Rand, D. G., J. D. Greene, and M. A. Nowak. 2012. "Spontaneous giving and calculated greed". *Nature* 489:427-430.

Reips, U.-D. 2001. "The Web Experimental Psychology Lab: Five years of data collection on the Internet". *Behavior Research Methods, Instruments, and Computers* 33(2):201-11.

Schmelz, K. and A. Ziegelmeyer. 2012. "Reactions to control with(out) background control: Evidence from the internet and the laboratory". *mimeo. Working paper. Max Planck Institute of Economics.* `http://www.econ.mpg.de/files/2011/staff/schmelz/Background_Control_Schmelz_Ziegelmeyer.pdf`.

Schulte-Mecklenbeck, M. and O. Huber. 2003. "Information search in the laboratory and on the web: With or without an experimenter". *Behavior Research Methods, Instruments, and Computers* 2:227-235.

Shavit, T., D. Sonsino, D., and U. Benzion, U. 2001. "A comparative study of lotteries-evaluation in class and on the web". *Journal of Economic Psychology* 22:483-491.

Suri, S. and D. J. Watts. 2011. "Cooperation and contagion in web-based, networked public goods experiments". *PLOS ONE* 6:e16836.

von Gaudecker, H.-M., A. van Soest, and E. Wengstroem. 2012. "Experts in experiments: How selection matters for estimated distributions of risk preferences". *Journal of risk and uncertainty* 45:159-190.