

Jahrbuch für Handlungs- und Entscheidungstheorie

Marc Debus · Markus Tepe
Jan Sauermann *Hrsg.*

Jahrbuch für Handlungs- und Entscheidungstheorie

Band 11



Springer VS

Jahrbuch für Handlungs- und Entscheidungstheorie

Reihe herausgegeben vom Sprecherteam des Arbeitskreises Handlungs- und Entscheidungstheorie der Deutschen Vereinigung für Politikwissenschaft

Weitere Bände in der Reihe <http://www.springer.com/series/12723>

Marc Debus · Markus Tepe ·
Jan Sauermann
(Hrsg.)

Jahrbuch für Handlungs- und Entscheidungstheorie

Band 11



Hrsg.

Marc Debus
Fakultät für Sozialwissenschaften
Universität Mannheim
Mannheim, Deutschland

Jan Sauermann
Center for Comparative Politics
Universität zu Köln
Köln, Deutschland

Markus Tepe
Sozialwissenschaften
Universität Oldenburg
Oldenburg, Deutschland

ISSN 2627-3268 ISSN 2627-3276 (electronic)
Jahrbuch für Handlungs- und Entscheidungstheorie
ISBN 978-3-658-23996-1 ISBN 978-3-658-23997-8 (eBook)
<https://doi.org/10.1007/978-3-658-23997-8>

Die Deutsche Nationalbibliothek verzeichnetet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer VS

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2019
Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.
Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.
Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer VS ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature
Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Editorial

Mit dem vorliegenden elften Band des Jahrbuchs für Handlungs- und Entscheidungstheorie ist es gelungen, die seit 2001 bestehende Schriftenreihe fortzusetzen und die themenoffene Struktur des Jahrbuchs beizubehalten, um weiterhin ein attraktives Forum für das mittlerweile breite Spektrum handlungs- und entscheidungstheoretischer Beiträge zu bieten. Vor diesem Hintergrund freut es uns sehr, in dieser Ausgabe drei originäre Forschungsaufsätze und vier Kurzanalysen präsentieren zu können. Es sei angemerkt, dass fünf der sieben Beiträge in englischer Sprache verfasst sind, was die zunehmende Internationalisierung der deutschen Politikwissenschaft verdeutlicht und auch die Entscheidung, das Jahrbuch bilingual zu gestalten, mehr als unterstützt.

Die in der diesjährigen Ausgabe des Jahrbuchs publizierten Beiträge decken einen weiten inhaltlichen Themenbereich ab, der die Fruchtbarkeit von Theorien rationalen Handelns und Entscheidens beeindruckend widerspiegelt. Auch im Hinblick auf die verwendeten Methoden weisen die Beiträge des vorliegenden Jahrbuchs eine große Vielfalt auf.

Insgesamt vier Beiträge befassen sich mit räumlichen Modellen der Politik. Der Aufsatz von Michael Jankowski und Kamil Marcinkiewicz untersucht auf Basis des räumlichen Wählermodells die Determinanten legislativen Entscheidens. Auf Grundlage von Roll Call-Daten des tschechischen Parlaments wird der Zusammenhang zwischen dem parlamentarischen Abstimmungsverhalten und der Regierungsmehrheit (formal minority, substantive minority, caretaker government und majority government) analysiert. Der Beitrag von Guido Tiemann befasst sich mit der Modellierung von Risiko und Unsicherheit in räumlichen Wählermodellen und deren empirischer Validierung mit Hilfe der CSES-Daten. Seine Ergebnisse zeigen, dass Wähler sehr viel weniger risikoavers sind als vielfach theoretisch angenommen. Stattdessen tendieren sie zu einem risikoneutralen

Verhalten. Die Kurzanalyse von Thomas Bräuninger, Marc Debus, Jochen Müller und Christian Stecker berichtet die Ergebnisse einer Expertenumfrage zur programmatischen Ausrichtung der deutschen Parteien im Herbst 2017. Neben der Positionierung der etablierten Parteien erlauben die präsentierten Daten auch eine Einordnung der erstmals im Bundestag vertretenen Alternative für Deutschland. Ingrid Mauerer und Micha Schneider zeigen in ihrer Kurzanalyse am Beispiel des Themas Migration, dass sich die Wähler bei der Einstufung der Parteien häufig unsicher sind. Die Autoren stellen einen statistischen Ansatz zur Messung wahrgenommener Parteipositionen vor, mit dem sich die resultierenden verzerrenden Effekte aus den beobachteten Platzierungen herausrechnen lassen.

Die verbleibenden drei Beiträge, die in dieser Ausgabe des Jahrbuchs für Handlungs- und Entscheidungstheorie präsentiert werden, sind Belege für die in den vergangenen Jahren stark gestiegene Bedeutung experimenteller Methoden in der Politikwissenschaft. Zwei Studien befassen sich mit dem Einfluss von Fairness und Gerechtigkeit in Verhandlungen. Helena Hagauer, Bernhard Kittel und Manuel Schwaninger verwenden in ihrem Forschungsaufsatz das bislang vor allem in der Soziologie herangezogene Modell der Frame-Selektion und argumentieren, dass das Verhalten von der Wahrnehmung und Interpretation der Situation durch das Individuum geprägt wird. Die These wird im Rahmen von Laborexperimenten zu Verhandlungsspielen überprüft, in denen die Wahrnehmung der Situationen erhoben und mit dem tatsächlichen Verhalten der Probanden in Beziehung gesetzt wird. Jan Sauermann untersucht die Effekte von Reputationsbildung im Ultimatumsspiel. Die experimentellen Befunde seiner Kurzanalyse zeigen, dass Reputationsbildung zu faireren und effizienteren Verhandlungslösungen führt. Christine Prokop nutzt in ihrer Kurzanalyse eine Reihe laborexperimenteller Designs, um zu eruieren, ob sich anhand von subjektiven Risikoeinstellungen Risikoverhalten vorhersagen lässt und auf welche Weise dieser Zusammenhang durch das Geschlecht moderiert wird. Ihre Ergebnisse zeigen keinen generellen direkten Zusammenhang zwischen Einstellungen und Verhalten, jedoch einen deutlichen Einfluss des Geschlechts. So handeln weibliche Versuchspersonen deutlich risikoaverser als männliche Probanden.

Wir wollen das Editorial der elften Ausgabe des Jahrbuchs für Handlungs- und Entscheidungstheorie auch dazu nutzen, um an den 25. Geburtstag des Arbeitskreises zu erinnern. Der Arbeitskreis Handlungs- und Entscheidungstheorie wurde im Sommer 1993 mit dem Ziel gegründet, Rational Choice-Ansätze für die deutsche Politikwissenschaft fruchtbar zu machen. Zu letzterem hat der Arbeitskreis beigetragen: nicht nur auf Basis der elf Jahrbücher (einschließlich dieser Ausgabe) für Handlungs- und Entscheidungstheorie, sondern auch auf

Grundlage der seit 2005 jährlich stattfindenden Tagungen und den Aktivitäten auf den Kongressen der DVPW. Die jeweils dort präsentierten Beiträge sind nicht nur im Jahrbuch des Arbeitskreises erschienen. Vielmehr ist eine Reihe dieser Beiträge in nationalen wie internationalen Fachzeitschriften veröffentlicht worden. Wir möchten uns an dieser Stelle bei allen Teilnehmerinnen und Teilnehmern der Tagungen des Arbeitskreises bedanken sowie insbesondere bei den Kollegen, die seit der Gründung des Arbeitskreises als dessen Sprecher gewirkt und damit zu Etablierung und Erfolg der analytischen Politikwissenschaft in Deutschland beigetragen haben: André Bächtiger, Joachim Behnke, Thomas Bräuninger, Hans-Peter Burth, Ulrich Druwe, Bernhard Kittel, Volker Kunz, Eric Linhart, Thomas Plümper und Susumu Shikano.

Unser Dank gilt natürlich auch den Gutachterinnen und Gutachtern, die die für das Jahrbuch für Handlungs- und Entscheidungstheorie eingereichten Beiträge kritisch und konstruktiv evaluieren. Im Editorial der ersten, von Ulrich Druwe, Volker Kunz und Thomas Plümper im Jahr 2001 herausgegebenen Folge des Jahrbuchs für Handlungs- und Entscheidungstheorie werden neben den Qualitätssicherungsmaßnahmen, maßgeblich determiniert über den Peer Review-Prozess, die interdisziplinäre und internationale Orientierung sowie die Mikrofundierung von Erklärungsansätzen, die die Analyse von politischen und sozialen Phänomenen auf Makroebene unterfüttern, als zentrale Eigenschaften der Beiträge im Jahrbuch genannt. Die bisher erschienenen Jahrbücher sind – aus unserer Sicht – diesen Vorgaben mehr als verbunden geblieben und wir hoffen, dass dies auch für den elften Band des Jahrbuchs gilt.

Mannheim, Oldenburg und Köln

Die Herausgeber

Inhaltsverzeichnis

Teil I Aufsätze

Fairness in Bargaining: How Self-selected Frames Affect the Fairness of Negotiation Outcomes	3
Helena Hagauer, Bernhard Kittel and Manuel Schwaninger	
Varieties of Legislative Voting Patterns.	33
Michael Jankowski and Kamil Marcinkiewicz	
Zur Spezifizierung von Risiko und Unsicherheit in räumlichen Modellen	63
Guido Tiemann	
Die programmatischen Positionen der deutschen Parteien zur Bundestagswahl 2017: Ergebnisse einer Expertenbefragung	93
Thomas Bräuninger, Marc Debus, Jochen Müller und Christian Stecker	

Teil II Kurzanalysen

Perceived Party Placements and Uncertainty on Immigration in the 2017 German Election	117
Ingrid Mauerer and Micha Schneider	
Risk Attitudes, Gender, and Risk Behavior: Evidence from Two Laboratory Experiments	145
Christine Prokop	
Reputation Formation in Bargaining. Partners vs. Strangers in the Ultimatum Game	179
Jan Sauermann	

Teil I

Aufsätze



Fairness in Bargaining: How Self-selected Frames Affect the Fairness of Negotiation Outcomes

Helena Hagauer, Bernhard Kittel and Manuel Schwaninger

Abstract

We investigate whether fair behavior in negotiations depends on the frame given by an actor to a situation. To test this proposition, we study secondary data from bargaining experiments. In the experiments, subjects were asked what they thought the situation was about. We compare this perception with their behavior in the negotiations. The results show that subjects with a prosocial frame were more likely to distribute the resource equally than those with a proself frame. Investigating the factors that lead to the adoption of a prosocial or proself frame, we find that minor situational differences do not influence the choice of a frame, whereas factors which can be traced back to socialization exert considerable influence on the selection of a frame.

This paper reports results of the research group FOR2014 “Needs-based justice and distribution procedures”, project B1 “Needs-based justice and distributive preferences in social exchange networks” (<http://bedarfsgerechtigkeit.hsu-hh.de/>). We gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, KI 1419/2-1) and the Austrian Wissenschaftsfonds (FWF, I1888-G11).

H. Hagauer · B. Kittel · M. Schwaninger (✉)
Institut für Wirtschaftssoziologie, Vienna, Austria
E-Mail: manuel.schwaninger@univie.ac.at

H. Hagauer
E-Mail: helena.hagauer@univie.ac.at

B. Kittel
E-Mail: bernhard.kittel@univie.ac.at

1 Introduction

Bargaining is usually associated with the idea of negotiating the best deal for oneself. According to the traditional economic assumption of self-interest, individuals are expected to strive for the highest possible payoff, regardless of their negotiating partners' outcomes. However, empirical and in particular experimental evidence shows that in many cases, people also take others' interests into account, distribute goods more equally and thus try to act in a fair manner.

But why do people behave in such a "fair" or "just" way when bargaining? What makes them align their behavior with standards of fairness or justice? The most common explanations for fair behavior in bargaining situations are altruism, reciprocity and avoidance of punishment or costs related to unfair behavior (Buchan et al. 2004, p. 183). Implicitly these explanations allege that the preferences for fair outcomes are individually stable and similar in any given situation. However, empirically the individual willingness to give up payoff for someone else can vary considerably. While people seek to maximize payoffs at the stock market, the same people may strive for a fair solution when they negotiate the time to meet close friends. The individual goals seem to depend on one's interpretation of the bargaining context. The essential precondition for acting in a fair way is the perception of fairness as a relevant orientation in a given situation. In other words, fair behavior depends on how individuals perceive the situation in which they are, which frame they attach to the situation, and, ultimately, what they think is the 'right' thing to do.

In the economic literature the idea that individuals interpret situations differently and that their behavior varies accordingly is far from commonplace. Nobel Laureate Vernon Smith holds in his induced value theory (Smith 1976) that experimental subjects' preferences can be controlled and that with the help of a reward system, the experimenter can "induce" new values according to which subjects will act. Individual differences between subjects become negligible in this process and subjects act mainly in accordance with the newly induced values. Yet, other economic studies have pointed to the fact that presenting an experimental situation in a different frame can have effects on subjects' behavior (Cookson 2000; Eriksson and Strimling 2014; Gerlach and Jaeger 2016; Kahneman and Tversky 1979; Liberman et al. 2004; Tversky and Kahneman 1981).

In this chapter, we go one step further by suggesting that even in a neutrally framed situation, actors' behavior varies considerably, depending on the way they perceive the situation. We are thus adopting a frame selection theory perspective (Esser 2010, 2018; Kroneberg 2005, 2007) on the interplay between the definition

of a situation and an actor's behavior. More specifically, we show that fair behavior in negotiations depends on the frame given by an actor to a situation. To investigate this relationship, we study secondary data from bargaining experiments. Experiments are especially suitable for this purpose because they confront subjects with a highly standardized, abstract situation. In order to choose an action, they first have to ask themselves the question, 'What is going on here? What is this experiment about?'. That is, they have to frame the situation. In the experiments, subjects were asked what they thought the experiment was about. We used this data and compared it with their behavior in the negotiations. The results show that the frame assigned by a participant to the situation explains much of the observed heterogeneity of behavior in experimental negotiations.

The chapter is structured as follows: In the following section, we outline the development of the related literature about the definition of the situation and frame selection theory and compare it with the concept of frames used in experimental economics, before sketching the importance of fairness frames in bargaining. In Sect. 3 we describe the experimental setup and develop our hypotheses about the relationship between the subjective definition of the situation in the game and actual behavior in the game. We present our results in Sect. 4 and discuss them in Sect. 5. The chapter closes with concluding statements.

2 Related Literature

The idea that individual actions depend on subjective perceptions of the situation looks back on a century-long history in sociology. In comparison, relatively recently economists also started to account for the way in which frames and the individual perception of a situation affect behavior, especially in the laboratory. In the following we review the two strands of literature and their relations. At the end of this section, we briefly explore the literature dealing with frames and fairness in a bargaining context.

2.1 Frames in Sociology

William I. Thomas, one of the most notable representatives of the Chicago School of Sociology, first introduced the concept of the *definition of the situation* in an essay about primary-group norms and their influence on the educational system. He argued that a society develops a system of norms and values or *code of behavior* which regulates how its members' behavior is evaluated. This code is developed in a process called the definition of the situation. Consequently,

depending on the definition of a situation, identical behavior may be considered “good or bad” (Thomas 1917, p. 167 f.). According to Thomas, this defining of the situation evolves in the course of socialization.

The concept of the definition of the situation was further developed and applied in another sociological classic, *The Polish Peasant in Europe and America* (Thomas and Znaniecki 1958), where the authors investigated how Polish immigrants adapted their behavior from the traditional surroundings in the Polish countryside to modern Chicago:

[T]he definition of the situation is a necessary preliminary to any act of the will, for in given conditions and with a given set of attitudes an indefinite plurality of actions is possible, and one definite action can appear only if these conditions are selected, interpreted and combined in a determined way and if a certain systematization of these attitudes is reached, so that one of them becomes predominant and subordinates the others (Thomas and Znaniecki 1958, p. 68).

A few years later, Robert Park and Ernest Burgess, two further central representatives of the Chicago School of Sociology, suggested that “every single act, and eventually all moral life, is dependent upon the definition of the situation. A definition of the situation precedes and limits any possible action, and a redefinition of the situation changes the character of the action” (Park and Burgess 1921, p. 764).

Today, Thomas’ concept of the definition of the situation is usually known by the name *Thomas theorem*: “If men define situations as real, they are real in their consequences” (Thomas and Thomas 1928, p. 572). This reading is at least partly due to Robert K. Merton, who singled it out as the essence of his own concept of the self-fulfilling prophecy (Egloff 2015, p. 21). A “self-fulfilling prophecy is, in the beginning, a false definition of the situation evoking a new behavior which makes the originally false conception come true” (Merton 1948, p. 195). As Thomas has emphasized and as we shall show in the experimental data, Merton’s self-fulfilling prophecy stresses the behavioral repercussions of defining situations in one or the other way.

Alongside Merton’s, we find various versions of the definition of the situation in key contributions to the sociological debate after WWII. For example, in Talcott Parsons’ concept of the *unit act*, the definition of the situation corresponds with the assumption that every act is founded on a normative orientation. This normative orientation determines the choice of action in two ways: On the one hand, it leads to a restriction in the number of available means of action, since not all means are normatively appropriate. On the other hand, the normative orientation structures the goals of acting by narrowing down which states actors consider worth striving for (Kroneberg 2011, p. 64).

Another influential sociologist of the 20th century, Erving Goffman, opened his *Frame Analysis: An Essay on the Organization of Experience* with a reference to the *Thomas theorem*, remarking that “[p]resumably, a ‘definition of the situation’ is almost always to be found, but those who are in the situation ordinarily do not *create* this definition, even though their society often can be said to do so; ordinarily, all they do is to assess correctly what the situation ought to be for them and then act accordingly” (Goffman 1986, p. 1 f.). For Goffman, the essential question individuals are confronted with in every situation is “What is it that’s going on here?” (Goffman 1986). Following Gregory Bateson, Goffman uses the term *frame* to describe categories used by actors to define situations. A frame then determines what is perceived to constitute a situation and in what way. Frames thus have a selective function, they are at the same time exclusive and inclusive. This selective function refers both to the way an actor experiences the situation and to her behavior (Kroneberg 2011).

The definition of a situation is of such importance for an individual’s behavior that it has been integrated into many sociological theories. Also scholars influenced by rational choice theory have taken up on it, using the concept as a basis for a comprehensive theory of action. For example, Siegwart Lindenberg’s theory of goal-framing sets out to explain an individual’s actions with recourse to the relative relevance of conflicting goals. According to Lindenberg, in each situation humans are guided by the goal that is activated at the moment. This so-called foreground goal subsequently dictates how an individual perceives her environment and which actions she takes. While other goals may still be active in the background, the overarching goal constitutes the definition of the situation. Lindenberg identifies three overarching goals or “goal-frames”: the *hedonic goal* aims at improving the way an actor is feeling at the moment, the *gain goal* implies striving for financial or economic gains and the *normative goal* accounts for the fact that individuals wish to act appropriately by obeying social expectations. *A priori*, the *hedonic goal* is strongest, making it the least dependent on social context. The *gain goal* needs institutions guaranteeing future revenues, while the *normative goal* is influenced even more by the social context in the form of institutions, moral and sanctions (Lindenberg 2013).

One of the major criticisms of Lindenberg’s goal-framing theory has been the absence of an explanation why a certain goal-frame is activated (Liebig and Sauer 2013). This and other problems are addressed by Hartmut Esser. His model of frame selection (MFS), which was refined and formalized by Clemens Kroneberg, sets out to combine elements of the utilitarianistic, normative and interpretative paradigms into a general theory of action (Esser 2010). The principal mechanisms put forward in this model are *frame selection*, *script selection* and *action selection*.

Frame selection is the process whereby an individual interprets the situation by selecting a frame out of a set of cognitively available frames. Following the selection of a frame, actors may activate a script, i.e. a certain program of action such as norms, routines or other cultural schemes of reaction. Finally, an actor has to decide whether she follows the script or opts for an alternative action (Kroneberg 2007, p. 217).

Integrating the assumption of variable rationality, the model describes the conditions under which the frames are chosen in either a reflecting-calculating or an automatic-spontaneous mode. While in the reflecting-calculating mode, actors attempt to maximize their subjectively expected utility, in the automatic-spontaneous mode the choice of a specific frame depends on i) the mental anchoring of the frame, ii) the availability of objects significant for the frame in the current situation and iii) the strength of the connection between the objects and the frame (Kroneberg 2005). The MFS has been used to explain diverse phenomena, including childlessness (Eckhard 2014), trust (Rompf 2015), electoral participation and the rescue of Jews in WWII (Kroneberg et al. 2010).

2.2 Frames in Experimental Economics

Apart from the contributions by Ellingsen et al. (2012) and Eriksson and Strimling (2014) the concept of frames in behavioral and experimental economics has developed without taking notice of the sociological contributions to the topic. In an overview article about the relevance of framing effects in experimental practice by Gerlach and Jaeger (2016), for instance, none of the concepts discussed in the previous section are mentioned. “Framing effects” in an experimental economics context are said to be present when actors’ decisions vary according to the way experimental instructions are presented (Cookson 2000, p. 55). This understanding is very similar to the one used in political science, where framing effects are referred to as (minor) changes in the presentation of information that can lead to changes in opinion (Chong and Druckman 2007). There are different forms of framing, valence framing and context framing being the most discussed ones.

Valence framing is present when an experimental situation is depicted in an either positive or negative light. Ground-breaking work on valence framing was done by Kahneman and Tversky, who challenged the traditional expected utility stance according to which actors choose actions merely on the basis of expected utility. They showed that individuals are more willing to take risks when it comes to avoiding losses than when given the opportunity to make gains (Kahneman and Tversky 1979). In another well-known study, Tversky and Kahneman (1981)

confronted participants with two different programs to combat a fictitious “Asian disease”. The programs had the same expected value, but one implied a risky outcome while the other yielded a sure outcome. Individuals to whom the task was presented in terms of lives saved mostly opted for the sure outcomes whereas those to whom it was presented in terms of lives lost predominantly chose the risky option. Valence framing is usually explained with the help of prospect theory, according to which gains and losses have different value functions (Gerlach and Jaeger 2016).

Context framing on the other hand includes attaching different associations to the game and highlighting certain aspects of an experiment (Gerlach and Jaeger 2016). A frequently cited example for the former is Liberman et al. (2004). Labelling a Prisoner’s Dilemma as “The Community Game” led to significantly higher cooperation rates than introducing the same experiment as “The Wallstreet Game”. The second type of context framing is active in Brañas-Garza’s (2007) manipulation of the dictator game, where adding the sentence “Note that your recipient relies on you” to the instructions led to more generous behavior.

So far, we have discussed examples where an experimenter intentionally manipulated the instructions of a game in order to impose a certain frame on the participants. More recent research has revealed that even when instructions are formulated in a “neutral” way, individuals’ perceptions of the situation vary, leading to different behavior. For instance, Eriksson and Strimling (2014) compared contributions in a public goods game labelled “The Team Work Game” with contributions of participants who had received unlabeled instructions and spontaneously associated the experiment with teamwork and those who did not. The level of contributions was at the same level in both groups, while those who had received unlabeled instructions and not associated the game with teamwork contributed considerably less, suggesting that spontaneous associations and label framing exert similar effects in public goods games.

In the same vein, sociologist Yamagishi et al. (2013) conducted various experiments after which they asked participants to evaluate how much the respective game had resembled different real-life situations. They found a correlation between game behavior and the perception of the experiment as being similar to a collaborative situation. Moreover, the effect of a person’s game perception and of her social value orientation proved to be independent of each other, fending off possible objections that pro-social individuals automatically have pro-social views of the game, which, in turn, leads to their pro-social behavior. Instead, the effects of a person’s social value orientation and their perception of the game were found to have an independent and approximately equal impact on game behavior (Yamagishi et al. 2013).

The insight that people's individual perceptions (or definitions) of a game situation influence their behavior in the experiment should not come as a surprise since they are in line with the abovementioned concepts of the definition of the situation and frame selection. We contribute to this line of research by investigating the interplay between spontaneous associations and game behavior in a bargaining context.

2.3 Fairness Frames in Bargaining

Bargaining situations are ubiquitous in human life. From businesspeople negotiating a deal, to individuals buying a house, to families trying to settle where to go on holiday, bargaining forms an essential part of social interaction. Based on the aforementioned literature our study's aim is to elicit the role of frames in determining behavior in bargaining situations. More precisely, we are interested in the influence of frames on individuals' propensity to act in a fair manner in a bargaining context.

The idea that the salience of specific frames has an impact on justice or fairness attitudes in negotiations has not been discussed in the framing literature per se. Yet, various authors have mentioned that the perceived context might influence bargaining behavior. Max Weber in his *Economy and Society* already noted, “[w]here the market is allowed to follow its own autonomous tendencies, its participants do not look toward the persons of each other but only toward the commodity; there are no obligations of brotherliness or reverence, and none of those spontaneous human relations that are sustained by personal unions” (Weber 1978, p. 636). In other bargaining situations, individuals may act more fairness oriented. The key question is which value is salient in a certain situation and we can thus speak of value-related frames (Kroneberg 2009, p. 77). These frames then determine whether a person acts in accordance with a certain value, such as fairness.

Addressing individual preferences for specific principles of fairness or justice, Liebig and Sauer (2013) build on Lindenberg's goal framing theory. They claim that even within the normative frame, there exist equivocal standards, the choice of which is in close relation to the “logics of justice”. By the logics of justice, Liebig and Sauer mean the homology between the kind of social relationship and the respective principle of justice. Using Alan Fiske's four elementary forms of sociality (1992), they differentiate between four types of social relationships and their corresponding forms of justice. Thus, individuals who see themselves in a solidary community favor need-based justice, people in non-hierarchical networks and cooperatives are ruled by the principle of equality, while those in a

hierarchical relationship prefer entitlement-based justice. Finally, in short-term relationships among strangers, respectively in market interactions, the principle of equity is applied. One source of perceived injustice lies in the different perceptions of the kind of social relationship characterizing a certain context, or, to put it in the terminology used so far, in the different definition of the situation. Whether or not an individual obeys the logics of justice and applies the principle adequate in the given situation thus depends on the frame that has been activated (Liebig and Sauer 2013).

Investigating the fairness of managers' actions, Li et al. (2012) explicitly highlight the importance of frames for fairness considerations. They argue that managers will not include justice considerations in their decision making processes if "justice-related decision images" (Li et al. 2012) are not part of their decision frame. The decision frame is a result of either frame recognition, i.e. relating the current situation to a previously encountered situation and applying the same frame, or frame identification, i.e. coming up with a frame for a situation that has not been encountered thus far (Li et al. 2012).

In this section we discussed two potential deviations from the assumption underlying induced value theory that actors in an experimental situation act mainly due to incentives given in the situation itself. First, presenting the experimental situation in a different frame can lead to different behavior. Second, individuals come up with a frame for every situation they encounter and act according to their frame. Transferring these insights to our research question as to when people behave in a fair way in bargaining situations, we arrive at the general expectation that individuals are more prone to behave in a fair manner in bargaining situations if they frame the situation accordingly and that their choice of a frame is determined by situational factors and socialization. In the following section, we outline the experimental setup used to test these assumptions empirically.

3 Experimental Data

To test the propositions that we derived from frame selection theory we study a series of laboratory bargaining experiments including a total sample of 663 participants in 31 experimental sessions (Kittel et al. 2017; Schwaninger et al. 2018). The experiments were originally designed to examine different research questions, but their design makes them very suitable for our analysis: They all share the same basic bargaining design, they are neutrally framed and, what is crucial, after the negotiations the subjects were asked what they thought the situation was

about.¹ We categorize the open answers qualitatively into a *prosocial*, a *proself*, and a *neutral* frame and then compare them with the behavior in the negotiations. In sum, we stepwise vary four parameters of the negotiations in a between-subject design.² In the following we describe the basic bargaining design and the specific treatment differences.

3.1 Basic Bargaining Design

In all sessions, subjects were allocated into groups of three and negotiated in negatively connected three-line networks. A *three-line network* consists of a group of three subjects who bilaterally bargain how to divide a fixed resource among the network members. In a three-line network, one *central* subject is connected with two peripheral subjects, whereas the *peripheral* subjects are only connected with the one central subject. In a negatively connected network, each player can conclude only one contract, which implies in a three-line network that the central subject can only settle on one agreement with one of the peripheral subjects and cannot settle on two agreements separately. In other words, the central subject is the only one being able to form the necessary majority to make a decision. Thus, in negative three-line networks one peripheral subject is excluded from the decision and has at best only indirect influence on the negotiation outcome. Since the peripheral subjects are threatened by exclusion, they are expected to compete with each other for an agreement with the central subject. Therefore, the structure of the network differentiates between two distinctly different power positions, and consequently, central subjects are expected to obtain a larger share of the resource.

In all treatments the subjects had three minutes to arrive at a decision. Within these three minutes they were free to send as many numerical offers as they wanted, costless and unrestricted. Negotiations took place via the computer program, ruling out independent effects of face-to-face interaction. If the group could not agree on

¹Note that asking the participants about their perceived frame after the experiment is controversial. We will discuss the advantages and disadvantages of this procedure in more detail in the discussion section of the chapter.

²The language and visual set-up of the negotiations are almost identical across all experimental sessions. However, similar to other meta-studies, some collected control variables vary between treatments. We describe the differences later in this section.

any allocation, all three network members received no payoff. After all negotiations ended, a new round began and subjects were randomly reassigned into new groups of three subjects. Depending on the experiment, the subjects negotiated between five and 20 rounds, but by focussing only on the first five periods we can disregard this design difference. At the end of the experiments, one or more rounds were randomly selected, the payoffs of the chosen rounds converted into Euros and then paid to the participants. All participants were fully informed about the payment scheme.

One specific feature of the experimental design is that the agreeing dyad is able to divide the payoff not only between themselves bilaterally, but they are also able to include the third network member into their distribution offers. The authors call this *inclusive exchange*. All negotiations we study focus solely on inclusive exchange. Since common network exchange experiments are studied within an economic exchange framework, i.e. within a market frame, other studies (not reported here) restrict allocation offers on a bilateral division of payoffs by design, i.e. *exclusive exchange* (Molm 2014).

3.2 Treatments

Although the basic set-up is the same, four dimensions varied between the experiments. In what we refer to as treatment 1 (108 participants), subjects participated in exclusive exchange before or after engaging in inclusive exchange.³ It is the only treatment where participants rotated across network positions. While all subjects were randomly reallocated to new groups in each new round in all treatments, only in the first treatment it was possible to switch from a central to a peripheral position or vice versa. The set-up of treatment 2 (75 participants) differs from treatment 1 in that subjects only participate in inclusive exchange and in that positions are stable across all rounds. In the remaining treatments, the network position in round one determined the network position in all subsequent bargaining rounds.

³In this study, we concentrate on inclusive exchange and do not analyze data from the rounds with exclusive exchange. Still, the inclusive instructions might trigger a selfish frame. Yet, it is equally likely that the new opportunity to share the payoff across all three group members triggers a prosocial frame. Clear predictions are not possible in this case. Empirically, there are no differences.

Treatment 3 (192 participants) mirrors treatment 2, with the difference that the authors introduced needs in treatment 3. The needs were implemented as different threshold levels of the number of points to be reached in the negotiations in order to earn additional income in a subsequent real-effort task. The real-effort task included trivia questions, counting capital letters, IQ questions and simple math tasks. If the threshold was reached, the subjects were paid for every task that they solved correctly. However, the threshold levels were private information and no subject knew about the threshold level of the other network members. In treatment 4 (288 participants) this changed and in contrast to treatment 3, subjects had full information about all threshold levels in the network.

3.3 Individual Attributes

Apart from the negotiation outcomes, the authors of the experiments collected several additional variables to account for individual characteristics. We use these individual attributes for our analysis of socialization effects on frame and script selection, as will be outlined in the Hypotheses section. The individual characteristics were elicited in a post-treatment questionnaire including several socio-demographic questions. The available data allows us to control for gender, age, the experience in economic experiments, the field of study, the semester of their studies, and the mother tongue. Moreover, in 22 out of the 31 sessions we can use a measure of self-reported risk aversion. In the other 9 sessions the participants instead completed an incentivized task to approximate risk aversion (Holt and Laury 2002) prior to the negotiations.

Additionally, after answering how they perceived the game, the participants had to rank statements concerning the importance of different factors for the agreements. We can categorize the answers into *proself* and *prosocial*, whereby *proself* indicates that maximizing own payoffs was more important than other goals to this person. Therefore, we also have a self-reported measure of distributional goals of the participants.

3.4 Hypotheses

Table 1 sums up the different treatments and Fig. 1 shows the expected causal relationship between a situation and the consequential behavior. The first set of hypotheses describes the relationship between the given situation and the perceived frame. The second set of hypotheses then describes the relation between the frame and the behavior.

Table 1 Treatments and differences

	<i>I</i>	<i>2</i>	<i>3</i>	<i>4</i>
Treatments				
<i>Positions</i> ^a	Rotating	Stable	Stable	Stable
<i>Needs</i> ^b	No	No	Yes	Yes
<i>Needs Information</i> ^c	No	No	No	Yes
Dependent variable				
<i>Prop. new wording</i> ^d	0.50	0.36	0.50	0.50
<i>Observations</i>	108	75	192	288
Control variables				
<i>Risk measure</i>	Incentivized	Incent./s.-r.	Self-reported	Self-reported
<i>Socio-demographic info</i>	Yes	Yes	Yes	Yes

^aStable positions imply that subjects were either in a central or peripheral position throughout the whole experiment, while rotating means that they were sometimes in a central and sometimes in a peripheral position

^bIf needs were present, they were introduced in the form of a threshold that individuals had to reach in order to earn points in a subsequent real effort task

^cNeeds information refers to the question whether subjects were informed about other subjects' need thresholds

^dProportion of subjects who received a new version of the question as to what the experiment was about. (In the first sessions, the participants were asked: "If you had to describe the experiment in one word, what would it be?" However, this question also yielded an undesirably high amount of answers like "interesting" or "boring", which could not be categorized into prosocial or proselytizing frames. Therefore, in the remaining sessions the subjects were asked instead: "In one word, how did you experience the game?" before being asked, inspired by Goffman (1986), "In one word, in your opinion, what was this game about?". The latter procedure improved the quality of the data. The table shows the proportion of cases when the latter question was implemented. We control for the different wording of the question in our regression analysis.)

**Fig. 1** Causal relationship between situation and behavior

An experiment has the advantage to control the situation to a high degree. To derive our first set of hypotheses we utilize the features of an experiment in two ways: On the one hand, the number of possible frames and actions is limited in an experiment. Given that the participants learn that their only task in the laboratory bargaining experiment is to distribute payoff, there is little, although not no, room for individual interpretation of the situation. Essentially, a participant can either try to distribute the payoff equally or to maximize own payoffs. Thus, we can subsume the frames under two main categories. The *proself* frame indicates that a participant believes the experiment to be about making as much money as possible for him- or herself. This frame describes the situation as being similar to a market, rather competitive and individualistic. The other, *prosocial*, frame suggests that the participant interprets the situation as being about fairness and equality. This frame activates the idea of a community, associated with cooperation, fairness or solidarity.

On the other hand, the researchers control and systematically vary the exogenously imposed situation in a laboratory experiment. As Kroneberg (2005) notes, the availability of objects in the actual situation affects the choice of a frame. Since the situation is the same for all participants, the only situational differences that can causally affect the perception of the situation on average are the differences between treatments. Thus, the question is, how do the treatment differences affect the perception of the situation? We propose the following set of hypotheses:

H1a *Individuals in a powerful position are more likely to adopt a proself frame than those in a weak position or those who rotate across positions.*

H1b *Introducing need thresholds and making them salient for all group members increases the likelihood that individuals adopt a prosocial frame.*

The first hypothesis is in line with one of the basic observations in social exchange research according to which more powerful actors have a tendency to exploit this power in negotiations by maximizing their own profits (Cook and Emerson 1978; Cook and Gillmore 1984; Emerson 1972; Molm 2014; Molm et al. 2003). We also expect that knowing that one can rotate across both positions makes subjects more likely to adopt a prosocial frame, since motives such as reciprocity and social insurance become salient in this situation.

The second hypothesis follows Liebig and Sauer (2013). If close communities are more likely to distribute according to a need-based justice principle, vice versa, the notion of needs in the group should be more likely to trigger a communal and prosocial frame. However, we expect that knowing the others' thresholds is crucial and only one's own threshold level is not sufficient. Even though the notion of

needs might trigger prosocial frames in this situation as well, predictions are not clear, since knowing only the own needs might also highlight the importance to earn a lot of payoff for oneself.

Situational factors account only partially for the choice of a specific frame. The availability of frames and scripts mainly depends on a person's socialization (Kroneberg 2007, p. 217). We therefore use the available socio-demographic data which are indicative of differences in socialization to develop assumptions about the probability for choosing one or the other frame. For example, since values like altruism or empathy play a more important role in girls' upbringing than in boys' (Kane 2012; Rendtorff 2006), one may expect women to be more likely to adopt a prosocial frame than men. In particular, we expect that economics and business students are more likely to apply prosocial frames, since students who are concerned with profit maximization tend to self-select into these majors (Frey and Meier 2003). Frame selection theory also explicitly highlights that experience, in this case experience in experiments, changes the salience and availability of scripts in a situation such as the norms, conventions, routines, heuristics or emotional patterns (Kroneberg 2005). According to the dual process theory integrated in frame selection theory, participants with less experience in experiments should be more likely to adopt a reflecting-calculating mode than an automatic-spontaneous one. We are able to test whether and how individual attributes influence frame selection and behavior by including them into the regression analysis.

The second set of hypotheses deals with the relationship between selected frame and behavior in the experiment. One of the key assumptions of the model of frame selection is that "behavior is guided by frames and scripts" (Kroneberg et al. 2010, p. 6). The selection of a frame structures a person's view on the situation by providing a certain relevance structure and thereby simplifying the situation. Following the choice of a frame, a script (such as a norm, routine or other behavioral programs) may be activated. Finally, an actor needs to decide which action to take (Esser 2010, p. 55). The choice of a frame therefore has an influence on an individual's behavior by reducing the number of possible interpretations of a situation and by possibly activating a behavioral program. We distinguish between prosocial and prosocial frames. As far as behavior is concerned, we analyze the degree to which subjects approve of more equal allocations. We hypothesize the following:

H2a *Individuals in a prosocial frame have a lower likelihood than others to support more equal allocations.*

H2b *Individuals in a prosocial frame have a higher likelihood than others to support more equal allocations.*

4 Results

The experimental sessions were conducted at the Vienna Center for Experimental Economics and the WISO-laboratory at the University of Hamburg between April 2016 and March 2018 in German.⁴ The students' average earnings in the experiments were between 12 and 15 euros per hour. In this section we first analyze the effects of treatments and personal characteristics on the indicated frames. Then we report the relationship between frames, goals and observed behavior.

4.1 Participants' Definition of the Situation

The coding procedure used to categorize the subjects' frames involved three steps. First, as mentioned above, subjects were asked to describe in one word what they thought the experiment was about. This question yielded a plethora of different answers, ranging from "profit maximization", through "solidarity", to "speed". In a second step, we screened the data for similar answers, inductively developing categories. In the third step, two researchers independently assigned all associations into one of the categories. Comparing the results of the classification, 92% of the codings were identical after the first run. In the second run the two researchers independently revisited the other 8% of the data again. For any mismatches after the second run, we consulted a third researcher, and after some discussion we agreed on all classifications (Mayring 2000, 2015). Since some of the categories contained only very few examples, making them unsuitable for statistical analyses, we grouped the categories into thematic clusters. The three resulting root categories are *proself*, *prosocial* and *neutral*. The *proself* root category contains interpretations in connection with profit maximization, competition, power and bargaining. The *prosocial* root category subsumes associations with fairness and cooperation. Answers that could neither be categorized as *proself* nor as *prosocial* were put into the *neutral* category.⁵ Many of the statements in this category refer to the speed of the game or emotions felt during the experiment; others concern the difficulty of the experiment or are very general associations. Table 2 reports the root categories, subcategories and examples for each.

⁴The authors used z-Tree (Fischbacher 2007) to program the experiment, in Vienna ORSEE (Greiner 2015) and in Hamburg hroot (Bock et al. 2014) to recruit the participants.

⁵Note that being in the neutral category does not exclude having a *proself* or *prosocial* stance, it only means that this stance did not become salient, which can be due to the way we asked for frames.

Table 2 Categories and typical examples

Root category	Subcategory	Examples
Proself	Profit maximization	<i>Profit, greedy, selfish, ME</i>
	Competition	<i>Competition, blitzkrieg</i>
	Power	<i>Power, hierarchy</i>
	Bargaining	<i>Bargaining, deals, negotiation</i>
Prosocial	Fair	<i>Fair, justice</i>
	Cooperation	<i>Cooperation, teamwork</i>
Neutral	Fast	<i>Speed, stress</i>
	Luck	<i>Luck, chance</i>
	Difficulty	<i>Difficult, tricky, complicated, challenging</i>
	Like	<i>Good, interesting, fun, KOOL</i>
	Dislike	<i>Nonsense, worthless, boring</i>
	General	<i>Experiment, distribution, game</i>
	Discard	<i>2, colorful dots, student</i>

Note that we translated the examples from German to English for this table.

Altogether, 25% of the answers were categorized as proself, 24% as prosocial and 51% as neutral. Table 3 shows how the distribution varies across treatments and positions. Interestingly, central players are not more likely to perceive the situation in a proself frame than peripheral players (test for equality of proportions, $p=0.54$). If anything, central players are more likely to indicate prosocial frames ($p=0.05$). Furthermore, subjects who alternate between the central and the peripheral positions are also not more likely to choose prosocial frames than subjects in central positions ($p=0.86$). On the contrary, rotating subjects are more likely to report proself frames than other subjects ($p<0.01$), because fewer answers are categorized in the neutral category ($p=0.01$). Also, transparent need claims do not lead to a more prosocial perception of the situation ($p=0.57$).

Overall, the different treatments do not appear to significantly affect the average distribution of perceptions. This suggests an important empirical result for frame selection theory: Frame selection does not seem to be sensitive to small variations of a situation. Across all treatments the subjects are confronted with

Table 3 Associations across treatments and positions

Treatment	Position	Prosself (%)	Association	
			Prosocial (%)	Neutral (%)
1	Rotating	37	23	40
2	Central	12	32	56
2	Peripheral	28	20	52
3	Central	17	34	48
3	Peripheral	21	21	58
4	Central	27	26	47
4	Peripheral	22	22	56
	Total	25	24	51

the same basic problem in slight variations. Hence, in order to find differences in the perception of the situation, one may need larger differences in the administered tasks. The negative result might also be an artifact due to the rather high abstraction of the categorization. However, analyses using a more fine-grained analysis of the stated frames into variations of prosself and prosocial (not reported here) do not indicate clear patterns.⁶

In addition, if there are changes in the perception in respect to treatments, then they tend to contradict our hypotheses. We might have drawn the wrong conclusions from the theory. But, at least to us, it seems more likely that factors which can be traced back to socialization have a more salient effect on the perception when the situations do not differ enough. Thus, in the following we directly compare whether participants selected a prosself or prosocial frame in a logit regression and add the available controls for gender, mother tongue, semester and field of study. Moreover, we control for experience in experiments, which can influence the availability of certain frames. Note that we excluded the observations with neutral frames from our analysis.

⁶Note that there were some differences within the neutral root category: E.g., in treatment 1, 45% of the neutral answers categorize as “fast” and 15% as “like”, whereas in treatment 4, only 16% fall under the subcategory “fast” and 25% of the answers in the neutral root category can be subsumed under the “like” subcategory. However, as these differences are not unambiguously informative for our research question, we do not further take them into account.

Table 4 reports the results. As before, the treatments tend to have no robust effects on the selection of prosocial or prosocial frames. However, other variables improve the explanatory power. First, gender is a significant predictor of the perception of the situation. Women are more likely to perceive the negotiations as a cooperative situation. In fact, the finding that women tend to act in a more cooperative way than men is widespread in experimental studies. Second, the mother tongue affects the perception of the game. German-speaking people are more likely to report a prosocial perception of the bargaining situation than subjects whose mother tongue is not German. This might be attributed to cultural differences and different socialization, different linguistic processing, different expectations of social desirability, or different selection effects to attend experiments. Third, the longer a person has been studying, the more likely this person is to select a prosocial frame. The number of semesters correlates significantly with participants' age, which is why we chose to include only one variable (Pearson correlation 0.45). We also tested whether the interaction between semester and economics studies affects the selection of certain frames, but this is not the case.

4.2 Frame and Action Selection

Finally, we arrive at the question whether a subject's frame explains the behavior in the experiment. As before, we differentiate between prosocial, prosocial and neutral frames. Behavior then is represented by the outcome of the negotiations. We measure inequality of the outcome as the final range of payoffs within the groups (Molm et al. 2000). The range captures interpersonal payoff differences,⁷ but also correlates highly with the standard deviation and variance of the payoffs. A smaller range indicates a more equal, "fair", distribution, while a larger range means that payoff is concentrated in single subjects.

We focus on the central, powerful players, whose preferences determine the outcome both theoretically and empirically (Schwaninger et al. 2018). Also, central subjects' average earnings in the negotiations are 1.5 times higher than those of peripheral subjects. Since the periods are not independent from each other, we examine the influence of the frame of the central subject on the *average* inequality of the first five negotiation periods.

⁷ $\max(x_i, x_j, x_k) - \min(x_i, x_j, x_k) = \frac{1}{2}(|x_i - x_j| + |x_i - x_k| + |x_j - x_k|)$

Table 4 Logit regression models. Frame selection and socialization

	Dependent variable: frame (proself or prosocial)			
	Model 1	Model 2	Model 3	Model 4
Treatment 1 ^a (ref. treatment 2)	−0.527 (0.424)	−0.771 (0.461)	−1.087* (0.500)	−1.073* (0.500)
Treatment 3 ^b	0.197 (0.401)	0.175 (0.403)	−0.326 (0.442)	−0.313 (0.441)
Treatment 4 ^c	−0.057 (0.379)	−0.079 (0.381)	−0.517 (0.421)	−0.502 (0.420)
Peripheral position ^d		−0.355 (0.261)	−0.457 (0.278)	−0.443 (0.278)
Female			0.716** (0.242)	0.704** (0.242)
Mother tongue German			0.876** (0.305)	0.867** (0.305)
Field: business or economics ^e			0.241 (0.321)	0.027 (0.453)
Attends the 7th semester (or higher)			0.786** (0.259)	0.703* (0.286)
Attends the 2nd experiment (or more)			−0.130 (0.298)	−0.140 (0.299)
Interaction: field and semester ^f				0.434 (0.641)
Constant	0.057 (0.338)	0.301 (0.384)	−0.673 (0.453)	−0.646 (0.454)
Observations	323	323	323	323
AIC	450.857	450.996	428.154	429.690
BIC	465.968	469.884	465.930	471.244
Log Likelihood	−221.429	−220.498	−204.077	−203.845

**p<0.01, *p<0.05

^aRotating positions.^bNeed thresholds (intransparent).^cNeed thresholds (transparent).^dSubjects in peripheral position can only negotiate with the central player, whereas subjects in central position can interact with both peripheral players.^eSubject studies business or economics.^fSubject studies business or economics in the 7th semester (or higher).

Basically, the range increases from 5.80 to 9.59 points on average if the frame of central subjects is not prosocial but prosself (Mann-Whitney test, $p < 0.01$).⁸ The correlation between frame and inequality of outcomes is 0.37 (Pearson corr.). This result supports the hypothesis that the selected frames affect the negotiation outcomes. In the following, we aim to further examine the mechanisms linking the frame and the outcome.

In a first step, we test whether the self-reported goals mediate the relation between the frame and the outcome. The subjects were asked to rank the factors that most influenced their decision in the negotiations. If they stated that maximizing their own payoff was most important to them, we categorize their goal as “proself”. In some sense, the stated individual preferences may be even more informative than the observable outcomes, since the outcome could also depend on the weaker negotiation partner. The prosself goal and inequality of negotiation outcomes correlate significantly (Pearson corr. 0.51, $p < 0.01$), as do the prosself frame and the prosself goal (Pearson corr. 0.22, $p < 0.01$). Indeed, the tests suggest that the goal mediates the influence of the frame on the bargaining outcome in a statistically significant way (Sobel test, $p = 0.01$).

According to frame selection theory, following the interpretation of a situation, subjects may “activate corresponding programs of behavior or *scripts* (...) among which a selection has to be made” (Kroneberg et al. 2010, p. 7). Similar to frame selection, script selection depends on socialization and experience and adds an explanation to the relationship between a situation and an observable outcome. Table 5 shows ordinary least square regressions explaining inequality as a function of the prosself or prosocial frame and our available controls.

Most notably, the impact of the frame is stable and statistically significant across various models. In contrast, the treatments, or marginally different situations, have no effect on the level of inequality.⁹ Also, risk aversion does not explain whether central subjects are more likely to agree on a relatively equal or unequal allocation of payoffs. Interestingly, the socio-demographic variables which before helped to explain the selection of a frame, do not help in explaining the negotiation outcome. The only variable that we measure and that makes a difference, apart from the selected frame, is the experience in experiments. People who have already

⁸The range is equal to 7.3 points if we could not clearly assign the central subject either to the prosself or prosocial frame, i.e. neutral frame.

⁹Note, however, that we did not expect the measured outcome variable to vary significantly between treatments.

Table 5 OLS regression models. Influence of the frame on the distribution of payoffs

	Dependent variable: range of payoffs in the group			
	Model 1	Model 2	Model 3	Model 4
Prosocial frame (ref. proselyt frame)	−3.786*** (0.788)	−3.402*** (0.815)	−3.158*** (0.870)	
Treatment 1 ^a (ref. treatment 2)		1.127 (1.621)	1.575 (1.674)	2.773 (1.711)
Treatment 3 ^b		−1.000 (1.682)	−0.527 (1.757)	0.132 (1.821)
Treatment 4 ^c		0.251 (1.617)	0.746 (1.680)	1.735 (1.728)
Female			−0.858 (0.802)	−1.341 (0.825)
Mother tongue German			0.287 (0.980)	−0.058 (1.016)
Field: business or economics ^d			−0.071 (1.062)	−0.294 (1.105)
Attends the 7th semester (or higher)			−0.617 (0.949)	−1.277 (0.971)
Attends the 2nd experiment (or more)			−2.428* (0.972)	−2.326* (1.013)
Risk averse ^e			−0.918 (0.872)	−1.647 (0.885)
Constant	9.586*** (0.556)	9.111*** (1.572)	9.915*** (1.640)	8.431*** (1.656)
Observations	151	151	151	151
R ²	0.134	0.156	0.207	0.133
Residual standard error	4.843	4.829	4.781	4.983
F-Statistic	23.070	6.760	3.658	2.393

*** p<0.001, ** p<0.01, * p<0.05

^aRotating positions.

^bNeed thresholds (intransparent).

^c Need thresholds (transparent).

^dSubject studies business or economics.

^eAn individual is classified as risk averse either by the measurement of Holt and Laury (2002), i.e. the individual switched to the sure option before risk neutrality, or when they stated they are risk averse on a 10 point Likert scale in the questionnaire. The results are robust if we include the risk measurements separately.

attended more than one experiment are more likely to share payoffs with others. This result suggests that while socialization matters more for frame selection, script selection depends more on the experience in a certain situation.

5 Discussion

Returning to our initial question of how self-selected frames affect the fairness of bargaining outcomes, the comparison of experimental subjects' frames, grouped into a prosocial, a prosocial and a neutral category, has revealed that subjects who chose a prosocial frame were more likely to support more equal allocations. Vice versa, individuals who chose a prosocial frame were less likely to endorse even allocations. This result stands in contrast to induced value theory, according to which preference orders can be induced in experimental subjects by using a proper reward medium, thereby making their innate characteristics largely irrelevant (Smith 1976). Instead, our results imply that a subject's individual perception of the situation has an influence on her or his behavior in the experiment.

Our hypotheses on frame selection have been rejected. Neither did being in a powerful position lead to more prosocial frames nor did the introduction of need thresholds lead to more prosocial frames. Instead, factors related to socialization, such as gender, culture, or age, have a much stronger effect on the choice of a frame than minor changes in a relatively neutrally framed experimental setup.

At this point, one might argue that if frames can essentially be traced back to socialization only, we could use less complex theories than MFS that account for social motives when analyzing individual behavior. However, we do not believe that our results justify this conclusion for two reasons. First, even though we did not observe clear patterns regarding the frequency of prosocial and neutral frames in the different treatments, we did observe differences in the choice of frames in the 'neutral' category. However, these differences were not investigated further since they are ambiguous on a theoretical level and not clearly informative for the outcome variable of interest. Second, even though factors related to socialization did have explanatory power when it came to choosing a frame, these factors were not predictive of the differences in the distribution of payoffs. The chosen frame, on the other hand, had a significant influence on the distribution of payoffs. It is therefore worth investigating more closely the process that leads to the adoption of one or another frame. According to Esser and Kroneberg, a frame is chosen when there is a match between the external object (characterized by "significant situational cues") and an individual's internal mental model (Esser 2010, 2018;

Kroneberg (2007, 2009). In our study, the situations differed only in small treatment variations which did not lead to significantly different frames on average. The treatment differences thus might not contain the relevant, “significant” aspects to affect the symbolic definition of the situation. Other differences might well induce markedly different frame choices. To explore other aspects of a situation that essentially influence the choice of a frame would be an important endeavor for future research.

As for methodological implications, our study adds to the discussion about context-free vs. in-context instructions. The common guideline, following Smith (1976), is to refrain from presenting experiments in a context. In line with Eriksson and Strimling (2014) and Engel and Rand (2014), we conclude that when subjects are not provided with a context or frame for an experiment, they will come up with their own frame and act accordingly. Thus, presenting instructions in a context might even increase control.

We also note several methodological controversies which are interesting for future research on frame selection theory. One possible objection that could be raised to the experimental design is the fact that subjects were asked for the frames after the experiment, which entails the possibility that the concrete experience during the experiment might have influenced the response to the question about frame selection rather than vice versa. Alternatively, one could ask for the frames upfront, after subjects have read the instructions. However, by raising the salience, this solution may lead to a self-fulfilling prophecy and potential perceptual changes in the course of the experiment cannot be accounted for. The latter problem could be solved by asking for the frames both before and after the experiment. Yet, given the possibility of a self-fulfilling prophecy, outcomes of the experiment might be affected by this variation in the design. Hence, the position of frame elicitation questions is itself an important research question for experimental work.

Moreover, the way people were asked for frames might not have yielded frames in the sense outlined in the literature discussed above, that is, frames as main goals becoming salient in a specific situation. For instance, if somebody stated that the experiment was about greediness, he or she might have taken this frame as an incentive to act in a way contrary to the frame, that is, in an especially fair way. This might explain the much higher correlation observed between subjects’ self-reported goals during the experiment and their subsequent behavior in the course of the experiment as compared to a lower, yet still highly significant

correlation between frames and behavior.¹⁰ For future research, we thus suggest adapting the questions and giving participants more space to elaborate on their interpretation of the experimental situation.

6 Conclusion

We have made an attempt to examine frame selection theory empirically by studying secondary data from bargaining experiments. We were firstly interested in whether or not subjects are more likely to behave in a fair way when they perceive the situation as being about fairness and cooperation. Secondly, we aimed to examine which determinants empirically influence the choice of a frame.

The results suggest that there is indeed a correlation between the frame chosen by an individual and the way she behaved in the experiment. Subjects who stated that the experiment was about fairness were more likely to distribute the resource equally than those who thought the experiment was about maximizing their own payoff. Investigating the factors that lead to the adoption of a prosocial or proselyf frame we have found that small situational differences do not influence the choice of a frame, whereas factors which can be traced back to socialization exert considerable influence on the selection of a frame.

From our perspective, this study makes three main contributions to the discussion about frames. First, we have found empirical support for frame selection theory. It seems that much of the heterogeneity of behavior in similar controlled lab experiments can be explained by different perceptions of the situation. The data suggest that socialization primarily affects frame selection, while experience with the situation is more likely to affect script selection.

Second, the results raise questions about the scope of induced value theory. Induced value theory laid the foundation for experimental economic research. Clearly, the incentive structure of a situation affects the behavior of people. However, this does not happen in a void. People try to understand a situation and this perception then affects their behavior as well. Therefore, if a researcher finds treatment effects due to small differences in a situation, it may often be that frames moderate the outcome. In particular, when the incentive structure of a situation does not change, frame selection theory can help explain otherwise inexplicable differences in behavior.

¹⁰See also, Fazio and Towles-Schwen (1999), Fishbein and Ajzen (2010), Opp (2015, 2017).

Third, combining the previous two points has interesting methodological implications for experimental research. A situation is always interpreted in a context and the individual interpretation of a situation can explain heterogeneous behavior. In addition, we know that frames can be manipulated by explicitly placing the study in a context. Sometimes experimental instructions can even be better understood if they involve non-artificial examples. Therefore, in some cases it may be advantageous to present experimental instructions with a frame. Especially if a researcher knows the particular field of interest of the research question, framing the experiment can reduce behavioral noise and, hence, increase control over the experiment.

We believe that our study adds to a fruitful discussion about the importance of frames both in experiments and in everyday contexts. Future research on frame selection theory should use primary data and could try to ask about the perception of the game in different experimental games and at different points in time during the experiment. It would also be interesting to see which factors of socialization affect frame-, script-, and action-selection the most, as our study is limited in this respect.

References

- Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. Hroot: Hamburg Registration and Organization Online Tool, *European Economic Review* 71: 117–20.
- Brañas-Garza, Pablo. 2007. Promoting Helping Behavior with Framing in Dictator Games, *Journal of Economic Psychology* 28 (4): 543–49.
- Buchan, Nancy R., Rachel T. A. Croson, and Eric J. Johnson. 2004. When Do Fair Beliefs Influence Bargaining Behavior? Experimental Bargaining in Japan and the United States, *Journal of Consumer Research* 31 (1): 181–90.
- Chong, Dennis, and James N. Druckman. 2007. Framing Theory, *Annual Review of Political Science* 10: 103–26.
- Cook, Karen S., and Richard Emerson. 1978. Power, Equity and Commitment in Exchange Networks, *American Sociological Review* 43 (5): 721–39.
- Cook, Karen S., and Mary R. Gillmore. 1984. Power, Dependence, and Coalitions. In *Advances in Group Processes, Volume 1*, edited by Edward J. Lawler, 27–58. Greenwich, CT: JAI Press.
- Cookson, Richard. 2000. Framing Effects in Public Goods Experiments, *Experimental Economics* 3 (1): 55–79.
- Eckhard, Jan. 2014. Theoretical Explanations of Increasing Childlessness - Divergent Approaches and the Integrating Potential of the Frame Selection Theory, *Comparative Populations Studies* 39 (1): 49–72.
- Egloff, Rainer. 2015. Definition of the Situation: History of the Concept. In *International Encyclopedia of the Social & Behavioral Sciences*, edited by James D. Wright, 19–23. Amsterdam: Elsevier.

- Ellingsen, Tore, Magnus Johannesson, Johanna Mollerstrom, and Sara Munkhammar. 2012. Social Framing Effects: Preferences or Beliefs?, *Games and Economic Behavior* 76 (1): 117–30.
- Emerson, Richard. 1972. Exchange Theory, Part II: Exchange Relations and Networks. In *Sociological Theories in Progress, Volume 2*, edited by Joseph Berger, Morris Zelditch Jr., and Bo Anderson, 58–87. Boston: Houghton Mifflin.
- Engel, Christoph, and David G. Rand. 2014. What Does ‘Clean’ Really Mean? The Implicit Framing of Decontextualized Experiments, *Economics Letters* 122 (3): 386–89.
- Eriksson, Kimmo, and Pontus Strimling. 2014. Spontaneous Associations and Label Framing Have Similar Effects in the Public Goods Game, *Judgment and Decision Making* 9 (5): 360–72.
- Esser, Hartmut. 2010. Das Modell der Frame-Selektion. Eine allgemeine Handlungstheorie für die Sozialwissenschaften?, *Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderheft 50, Soziologische Theorie kontrovers*: 45–62.
- Esser, Hartmut. 2018. Sanktionen, Reziprozität und die symbolische Konstruktion einer Kooperations-, „Gemeinschaft“, *Zeitschrift für Soziologie* 47 (1): 8–28.
- Fazio, Russel H., and Tamara Towles-Schwen. 1999. The MODE Model of Attitude-Behavior Processes. In *Dual Process Theories in Social Psychology*, edited by Shelly Chaiken and Yaacov Trope, 97–116. New York: Guilford.
- Fischbacher, Urs. 2007. Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, *Experimental Economics* 10 (2): 171–78.
- Fishbein, Martin, and Icek Ajzen. 2010. *Predicting and Changing Behavior: The Reasoned Action Approach*. New York: Psychology Press, Taylor & Francis Group.
- Fiske, Alan P. 1992. The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations, *Psychological Review* 99 (4): 689–723.
- Frey, Bruno S., and Stephan Meier. 2003. Are Political Economists Selfish and Indoctrinated ? Evidence from a Natural Experiment, *Economic Inquiry* 41 (3): 448–62.
- Gerlach, Philipp, and Bastian Jaeger. 2016. Another Frame, Another Game? Explaining Framing Effects in Economic Games. In *Proceedings of norms, actions, games*, edited by Astrid Hopfensitz and Emiliano Lori, 1–10. Toulouse: Institute for Advanced Studies.
- Goffman, Erving. 1986. *Frame Analysis. An Essay on the Organization of Experience*. Boston: Northeaster University Press.
- Greiner, Ben. 2015. Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE, *Journal of the Economic Science Association* 1(1): 114–25.
- Holt, Charles, and Susan Laury. 2002. Risk Aversion and Incentive Effects, *American Economic Review* 92(5): 1644–55.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk, *Econometrica* 47 (2): 263–92.
- Kane, Emily W. 2012. *The Gender Trap: Parents and the Pitfalls of Raising Boys and Girls*. New York: New York University.
- Kittel, Bernhard, Sabine Neuhofer, and Manuel Schwaninger. 2017. Need-Based Justice in Social Exchange Networks. *FOR2104 Working Papers* 2017–04.
- Kroneberg, Clemens. 2005. Die Definition der Situation und die variable Rationalität der Akteure. Ein allgemeines Modell des Handelns, *Zeitschrift für Soziologie* 34 (5): 344–63.
- Kroneberg, Clemens. 2007. Wertrationalität und das Modell der Frame-Selektion, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59 (2): 215–39.

- Kroneberg, Clemens. 2009. Das Modell der Frame-Selektion. Grundlagen und soziologische Anwendung einer integrativen Handlungstheorie. Universität Mannheim.
- Kroneberg, Clemens. 2011. *Die Erklärung sozialen Handelns*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kroneberg, Clemens, Meir Yaish, and Volker Stocké. 2010. Norms and Rationality in Electoral Participation and in the Rescue of Jews in WWII, *Rationality and Society* 22 (1): 3–36.
- Li, Jie, Suzanne S. Masterson, and Therese A. Sprinkle. 2012. Beyond the Eye of the Beholder. In *Research in Management: Perspectives on Justice and Trust in Organizations*, edited by Linda L. Neider and Chester Schriesheim, 139–63. Charlotte, NC: Information Age Publishing.
- Lberman, Varda, Steven M. Samuels, and Lee Ross. 2004. The Name of the Game: Predictive Power of Reputations Versus Situational Labels in Determining Prisoner's Dilemma Game Moves, *Personality & Social Psychology Bulletin* 30 (9): 1175–85.
- Liebig, Stefan, and Carsten Sauer. 2013. Soziologische Gerechtigkeitsanalyse. Überlegungen zur theoretischen Fundierung eines Forschungsfeldes, *Analyse & Kritik* 35 (2): 371–94.
- Lindenberg, Siegwart. 2013. Social Rationality, Self-Regulation, and Well-Being: The Regulatory Significance of Needs, Goals and the Self. In *The Handbook of Rational Choice Social Research*, edited by Rafael Wittek, Tom A.B. Snijders, and Victor Nee, 72–112. Stanford: Stanford University Press.
- Mayring, Philipp. 2000. Qualitative Content Analysis, *Forum: Qualitative Social Research* 1 (2).
- Mayring, Philipp. 2015. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim, Basel: Beltz.
- Merton, Robert K. 1948. The Self-Fulfilling Prophecy, *The Antioch Review* 8 (2): 193–210.
- Molm, Linda D. 2014. Experiments on Exchange Relations and Exchange Networks in Sociology: In *Laboratory Experiments in the Social Sciences*, edited by Murray Webster Jr. and Jane Sell, 199–224. New York: Elsevier.
- Molm, Linda D., Gretchen Peterson, and Nobuyuki Takahashi. 2003. In the Eye of the Beholder: Procedural Justice in Social Exchange, *American Sociological Review* 78 (1): 128–52.
- Molm, Linda D., Nobuyuki Takahashi, and Gretchen Peterson. 2000. Risk and Trust in Social Exchange: An Experimental Test of a Classical Proposition, *American Journal of Sociology* 105 (5): 1396–1427.
- Opp, Karl-Dieter. 2015. Norms. In *International Encyclopedia of Social and Behavioral Sciences*, edited by James D. Wright, 5–10. Oxford: Elsevier.
- Opp, Karl-Dieter. 2017. When Do People Follow Norms And When Do They Pursue Their Interests? In *Social Dilemmas, Institutions, and the Evolution of Cooperation*, edited by Ben Jann, and Wojtek Przepiorka, 119–42. Berlin, Boston: De Gruyter Oldenbourg.
- Park, Robert E., and Ernest Burgess (eds.). 1921. *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.
- Rendtorff, Barbara. 2006. *Erziehung und Geschlecht: Eine Einführung*. Stuttgart: Kohlhammer.
- Rompf, Stephan Alexander. 2015. *Trust and Rationality. An Integrative Framework for Trust Research*. Wiesbaden: Springer.

- Schwaninger, Manuel, Sabine Neuhofer, and Bernhard Kittel. 2018. Offers Beyond the Negotiating Dyad: Including the Excluded in a Network Exchange Experiment, *Social Science Research* 10.
- Smith, Vernon L. 1976. Experimental Economics: Induced Value Theory, *The American Economic Review* 66 (2): 274–79.
- Thomas, William I. 1917. The Persistence of Primary Group Norms in Present-Day Society and Their Influence in Our Educational System. In *Suggestion of Modern Science Concerning Education*, edited by Herbert S. Jennings, John G. Watson, Adolf Meyer, and William I. Thomas, 157–97. New York: The Macmillan Company.
- Thomas, William I., and Dorothy S. Thomas. 1928. *The Child in America: Behavior Problems and Programs*. New York: Alfred A. Knopf.
- Thomas, William I., and Florian Znaniecki. 1958. *The Polish Peasant in Europe and America*. New York: Dover Publications, Inc.
- Tversky, Amos, and Daniel Kahneman. 1981. The Framing of Decisions and the Psychology of Choice, *Science* 211 (4481): 453–58.
- Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. Berkeley: University of California Press.
- Yamagishi, Toshio, Nobuhiro Mifune, Yang Li, Mizuho Shinada, Hirofumi Hashimoto, Yutaka Horita, Arisa Miura, Keigo Inukai, Shigehito Tanida, Toko Kiyonari, Haruto Takagishi, and Dora Simunovic. 2013. Is Behavioral Pro-Sociality Game-Specific? Pro-Social Preference and Expectations of pro-Sociality, *Organizational Behavior and Human Decision Processes* 120 (2): 260–71.



Varieties of Legislative Voting Patterns

The Impact of Majority, Minority, and Caretaker Governments

Michael Jankowski and Kamil Marcinkiewicz

Abstract

We discuss how legislative voting patterns are affected by different types of governments. The analysis relies on a novel data set containing complete voting records of all members of the Czech Chamber of Deputies from 1996 to 2013. By employing spatial models, we analyze the positions of parties under formal minority, substantive minority, caretaker, and majority governments. In contrast to previous research, which has mainly focused on between country comparisons, our research design allows for a longitudinal analysis of voting patterns in a stable context limiting the number of potential confounders. We show that majority governments and formal minority governments are characterized by a dominance of government-opposition voting, which is in line with our theoretical expectations. In contrast, left-right differences become more relevant in times of substantive minority governments. We further specify this finding by offering a discussion of patterns observed under caretaker governments. We show that they share important characteristics with substantive minority governments. The results expand our understanding of legislative position taking and have implications for the analysis of roll-call votes in parliamentary democracies.

M. Jankowski (✉)

Fakultät für Bildungs- und Sozialwissenschaften, Oldenburg, Deutschland

E-Mail: michael.jankowski@uni-oldenburg.de

K. Marcinkiewicz

Fakultät für Wirtschafts- und Sozialwissenschaften, Hamburg, Deutschland

E-Mail: kamil-michal.marcinkiewicz@uni-hamburg.de

Keywords

Parties · Coalitions · Government types · Legislative voting

1 Introduction

Legislators are crucial actors in representative democracies. Their individual decisions determine the direction of lawmaking, turning legislative voting into an important element of democratic process and representation (Martin et al. 2014). Voting records can thus be a source of valuable information making it possible to understand better not only the policy outputs, but also the dynamics leading to their emergence. Yet, only in the last two decades has the study of legislative voting started to expand intensively beyond the US Congress (Poole and Rosenthal 1997) to include also the European Parliament (Hix et al. 2006, 2007), national-level legislatures outside the US (e.g. Godbout and Høyland 2011a, b; Spirling and McLean 2007; Lyons and Lacina 2009; see Hug 2013 for an overview), the sub-national level (Stecker 2011) and even the United Nations General Assembly (Voeten 2000).

In general, voting in legislatures can be conceptualized as shaped by two different principals. On the one hand, legislators are independent actors who can vote according to their own beliefs. On the other hand, this independence of legislators is often restricted due to strategic considerations in modern party politics and coalition agreements (Kam 2009, 2014; Christiansen and Pedersen 2014a). Governments and coalitions would be unlikely to persist in parliamentary democracies when party factions are not able to form stable voting blocs for a longer time period. As a consequence, high levels of vote unity have been observed in parliamentary democracies suggesting that legislators almost always follow the line of their party and deviating voting behavior occurs rarely (Sieberer 2006; Carey 2007). However, while this research suggests that the parliamentary voting behavior of individual legislators is driven by their party membership, less is known about the systematic formation of voting blocs between parties in parliamentary systems. That is, are ideologically contiguous parties more likely to vote similar in parliaments or is the position taken by parties in legislatures driven by strategic considerations, for example, by their status as government or opposition party? In this article, we address this research question.

More precisely, we analyze the position of parties based on their legislative voting record in the Czech Republic by employing spatial models of parliamentary voting, which have become “the workhorse theory of modern legislative studies” (Cox 2001, p. 189). These scaling methods estimate legislators’ ideal points based on their voting record in parliament (Poole 2005; Carroll and Poole 2014). In the US Congress, these ideal points can function as an accurate reflection of individual

legislator's position in a low-dimensional policy space (Poole and Rosenthal 1997). Such an interpretation of ideal points is, however, not always accurate outside the US context (Rosenthal and Voeten 2004; Spirling and McLean 2007). In parliamentary systems, ideal-points and dimensions revealed by these models can capture both, the government-opposition divide and the ideological position of parties or legislators (Bräuninger et al. 2016; Müller and Stecker 2014). Hence, a question arises about the conditions fostering specific voting patterns and the time at which these patterns change. This problem was addressed by Hix and Noury (2016), who recently provided a comprehensive comparative analysis of the voting behavior in several legislatures differing with regard to their institutional context (see also Coman 2015; Louwerse et al. 2017). They find that voting in legislatures is most often dominated by government-opposition dynamics and not by ideological positions such as the left-right placement of parties. The latter is only dominant under very specific institutional settings "namely, in presidential systems with coalition governments and in parliamentary systems with minority governments" (Hix and Noury 2016, p. 2).

This paper builds on the study by Hix and Noury (2016). However, instead of comparing the revealed ideal points¹ of deputies between various countries and at different time-points, we keep the overall context as stable as possible. In order to secure the comparability between the analyzed cases we have collected comprehensive voting records of Czech legislators. The Czech Republic offers an ideal setting for the analysis as it is characterized by comparatively frequent changes in the type of government (Hloušek and kopeček 2008). Government turnovers from majority to minority status (or vice versa), even within a single legislative period, are frequent occurrences. Our data covers the complete voting records of legislators since 1996 allowing for the analysis of 18 years of legislative voting. By relying on a longitudinal analysis of legislative voting patterns within a single parliament, we are able to exclude confounders which hinder the comparability of cases in cross-national research. We are therefore able to validate, specify and extend the findings reported in Hix and Noury (2016). In particular, the frequent change between majority and minority governments enables us to analyze directly

¹It is important to note that we refer to ideal points in a very broad sense. In a more narrow understanding of the term, an ideal point is the position of a political actor in a low-dimensional policy space. In contrast, we refer to an ideal point in this paper as the position estimated based on standard scaling techniques although we are fully aware of the fact that this position is heavily influenced by non-spatial factors such as the government-opposition divide and party discipline.

how and when the change of the institutional context affects parties' voting behavior in legislatures. Moreover, the wide variety of government-types observed in the Czech Republic allows to provide a more detailed picture of these effects for different types of minority governments and for caretaker governments. That is, we overcome the *majority vs minority* government dichotomy and instead differentiate between *majority, formal minority, substantive minority and caretaker* governments. Finally, the article provides the first comprehensive analysis of parliamentary voting patterns in the Czech Republic. So far, only a few studies analyzed the Czech Republic and relied on small subsets of the available voting records (Noury and Mielcova 2005; Lyons and Lacina 2009). Moreover, these studies estimate legislators' ideal-points for an entire legislative period. As we demonstrate in this article, such an approach might be misleading as voting patterns are not stable within legislative periods in particular when government types change.

Our empirical analysis demonstrates that government-opposition dynamics best explain legislative voting patterns when a majority or formal minority government exists. In contrast, left-right differences between parties become more relevant when a substantive minority or caretaker government is in power. We also find some exceptions to this picture. In particular, we find that severe government crisis can also impact legislative voting patterns although the type of government does not change.

The paper is structured as follows. First, we briefly discuss the theoretical debate on the relationship between the institutional context and voting in legislatures and formulate two hypotheses. We then describe the collected data and our methodological approach. In the following section we offer an interpretation of parties' ideal points estimated under different types of government. Finally, we reflect upon the scientific relevance of our core findings.

2 Types of Governments and Parliamentary Voting

In this article, we adopt a longitudinal perspective which makes it possible to compare legislative voting patterns between different units of analysis, while keeping the overall institutional context constant. This approach shares crucial characteristics with a case study about the Czech Republic presented in the analysis of Hix and Noury (2016). While the approach of Hix and Noury (2016) is particularly helpful to analyze the effects of different political systems (presidential vs parliamentary systems), we think that variation in the institutional context of a parliamentary system is best studied in a longitudinal design, where most of the other institutional factors can be held constant. We provide such an analysis

for the effect of different government-types in the context of the parliamentary system of the Czech Republic. This also implies that we are not analyzing the difference between parliamentary and presidential systems. Therefore, the section below explains our expectations about legislative voting patterns under different types of government in parliamentary systems.

2.1 Majority Governments

We start our discussion of government types with majority governments. These are characterized by consisting of one or more parties which hold a majority of seats in the parliament. Majority governments are thus able to form majorities for each vote and do not rely on the support of opposition parties as long as the governing parties vote cohesively. In these cases, standard assumptions about parliamentary voting are likely to be violated (Spirling and McLean 2007; Bräuninger et al. 2016), meaning that legislators do not necessarily vote for the option which is closest to their own position. Instead, under a majority government, parties and legislators are likely to act strategically. Government parties will vote together in order to secure their majority. Opposition parties, in contrast, have also strong incentives to vote together regardless of their ideological differences (Dewan and Spriling 2011; Kam 2014). This implies that not only government parties are likely to vote as a bloc, but also that opposition parties cast the same vote “strategically to defeat the government” (Spirling and McLean 2007, p. 8).² Therefore, majority governments are likely to produce government-opposition cleavages instead of left-right voting patterns (Müller and Stecker 2014).

2.2 Minority Governments

In contrast to majority governments, minority governments are supported by too few legislators to form a majority in parliament without the help of other legislators (Kalandrakis 2015, p. 309). Furthermore, it is assumed that all parties whose representatives hold minister positions in the cabinet belong to a minority

²It is important to note that we discuss the general patterns of parliamentary voting and that exceptions to this rule are likely, for example, due to rebels who tend to vote more sincerely (Spirling and McLean 2007) or due to “free votes” in which no means of party discipline are employed. Yet, these exceptions should not affect the overall positioning of parties.

government. It is, nevertheless, important to go beyond these general characteristics and to differentiate between two variations of minority governments. As described by (Strøm 1990, p. 62) we can distinguish between *formal* and *substantive* minority governments. A *formal minority government* describes governments which enjoy support of one or more opposition parties (or independent legislators) allowing them to obtain the parliamentary majority on a regular and reliable basis. In contrast, *substantive minority governments* have either no formal support of other parties or the support of other parties is still not sufficient to form a majority. One might think of substantive minority governments as “true” minority governments, while formal minority governments resemble majority governments. Although substantive minority governments occur more frequently in western democracies (Mitchell and Nyblade 2008), both forms of minority government are relevant for our study as their distinctive characteristics can help us explain the results we observe in the Czech Republic.

The common characteristic of minority governments is their inability to form a majority on their own. In order to obtain a majority of votes in parliaments, minority governments have to look for allies among other parties at each vote (Laver and Schofield 1992; Tsebelis 1995). This holds true especially for substantive minority governments, as they cannot rely on the ongoing support of other legislators. Under these circumstances, it is reasonable to expect that the minority government seeks a minimal connected winning-coalition at each vote (Axelrod 1970).³ Both the government and the cooperating opposition party (or individual MPs) benefit from the collaboration. The government parties have to adjust their positions only as long as it is necessary to obtain a majority, while the cooperating opposition parties have the opportunity to shape policy according to their own preferences. This trade-off has some limitations as highlighted by Strøm, since “the greater the opportunities for the parliamentary opposition to influence legislative policymaking, the lower the benefits of governing” (Strøm 1984, p. 212). However, this does not influence the assumption that cooperation is most likely between parties of close ideological proximity. Put differently, legislative voting patterns in parliaments under substantive minority governments are likely to reflect the ideological positions of parties.

For the case of formal minority governments, we expect that (informal) agreements with other parties or individual legislators result in similar patterns

³Such a coalition might be surplus. Therefore, parties might also look for minimal-winning coalition. However, as the ‘coalition’ only lasts for one vote we consider it more realistic that the ideological proximity is more relevant in this context.

of behavior as formal coalition contracts. We can then expect the parliamentary voting patterns under formal minority governments to resemble majority governments as both government types command the parliamentary majority on regular and relatively stable basis (Godbout and Høyland 2011a, p. 462).

The formal minority governments in the Czech Republic differ somewhat from the ideal type proposed by Strøm, which requires further explanation. The two Czech governments we refer to as formal minority cabinets were in fact minority coalitions, which controlled either two (Klaus 2) or one (Topolanek 2) MPs short of absolute majority (50% + 1) in the parliament. Both could pass the investiture vote and hence be confirmed by the parliament after the whole CSSD faction (Klaus 2) or two CSSD MPs (Topolanek 2) left the chamber. They hence did not control the majority of seats in the parliament, but were in fact also not really in minority (Topolanek 2) or lacked only one vote (Klaus 2). Due to controlling exactly or almost 50% of seats they still had much better chance of getting their motions passed by the parliament than the two substantive minority governments (Zeman and Topolanek 1), which had to build majorities on vote-by-vote basis.

The Klaus 2 government had an agreement with the CSSD concerning the investiture vote. The MPs of the CSSD left the chamber during the vote to reduce the size of the required absolute majority in exchange for the position of a speaker of the Chamber for the leader of the CSSD, Zeman. Under the second cabinet of Topolanek, the opposition did not control the majority of seats in the parliament and could not actively block the government. All in all the difference between the substantive and formal minority governments in the Czech case is a question of strength of parliamentary support rather than existence of opposition party providing the government stable and explicit support in the parliament (compare Christiansen and Pedersen 2014b, p. 943).

2.3 Caretaker Governments

Differently than minority governments, caretaker governments have attracted only little attention of political scientists (for a discussion of Czech caretaker governments see Hloušek and Kopeček 2014), maybe due to the assumption that caretaker governments “usually do not intend to undertake any serious policy making during their stay in office” (Müller-Rommel et al. 2004, p. 877). Yet, contested voting also takes place under caretaker governments and since we identify three cabinets in the Czech Republic belonging to this category, it is important to analyze their effects on representation. Caretaker governments are usually formed

for a relatively short period of time when the composition of parliament makes the appointment of a classical, partisan government impossible. Their main objective consists in managing the daily state affairs until the new parliament is elected. Strøm defines caretaker governments as “nonpartisan” (Strøm 1990, p. 7) as they are led by cabinet members who do not act as party representatives. However, even a caretaker government is usually elected by a parliamentary majority in the investiture vote. Yet, the coalition formed to elect the cabinet is not necessarily ideologically coherent. Therefore, majorities on specific bills can be formed in the absence of coalition constraints and each party can adjust its behavior based on its own policy preferences. Hence, the behavioral consequences of caretaker governments should be similar to those observed under the substantive minority governments.

To sum it up, a caretaker government, as we understand it, is primarily characterized by the fact that it is nonpartisan. Hence, it is somewhat similar to an expert government. Differently than some expert governments, however, the caretaker governments are formed to “keep the government functioning until the next duly constituted cabinet comes to power” (Conrad and Golder 2010, p. 125). Conrad and Golder (2010) differentiate between “new” caretaker governments and “continuation” caretaker governments. We focus on the former category. The “new” caretaker governments are expected to pass the investiture votes as opposite to “continuation” caretaker governments. In spite of its nonpartisan and transitory character two of three caretaker governments included in our study gained confidence of the majority of the members of parliament and existed until the new elections took place. The Rusnok government did not pass the investiture vote and was unable to fulfill its function. In spite of the fact that it consisted of nonpartisan experts it does not display characteristics of a caretaker government presented in Table 1. It did not gain support of the majority of MPs in the investiture vote and it was compelled to build “vote by vote” majorities. It must, hence, be categorized as a substantive minority government.

Table 1 Classification of government types based on majority in an investiture vote and vote by vote majority building

		“Vote by Vote” Majority building	
		Yes	No
Absolute Majority in investiture vote (Minimum 100 + 1 MPs)	Yes	Caretaker (LR)	Majority (GO)
	No	Substantive minority (LR)	Formal minority (GO)

2.4 Overview of Ideal Types and Formulation of Hypotheses

The four aforementioned types of governments can be classified with respect to two dichotomously coded theoretical dimensions. The first dimension displayed in the rows of Table 1 reflects whether it is necessary for the cabinet parties to create majorities on vote-by-vote basis. When it is sufficient to secure the absolute majority of votes consisting of at least 101 MPs (of 200 MPs) in an investiture vote we may speak of either majority or caretaker government. If, on the other hand, the ruling parties control fewer than 50% plus 1 parliamentary seats the government in question is either a substantive or a formal minority. The second dimension is more difficult to operationalize as assignment of a government to one of the two categories depends on existence (or lack) of a mechanism (formal or informal) making it possible to generate stable and reliable majorities in parliamentary votes. With respect to the second dimension we differentiate between majority and formal minority governments on the one hand, and substantive minority and caretaker governments on the other. While the former types can rely on mechanisms securing majority, the latter are forced to build majorities on a vote by vote basis (Hix and Noury 2016).

Obviously, the second of the two theoretical dimensions we use in our classification scheme has important observable consequences for parliamentary behavior of legislators as we expect them to be the most relevant factor for influencing legislative voting patterns. Hence, what matters most is not a formal minority or majority status based on the strength of cabinet parties, but whether a permanent parliamentary majority supporting the government exists. When the government is able to form a majority on a regular basis, then the opposition parties should be unlikely to support the government. Consequently, our first hypothesis reads as follows:

H1a: The government-opposition divide explains parliamentary voting patterns under formal minority governments and majority governments.

The broad agreement between many parties to build a caretaker government does not, on the contrary, lead to the emergence of the new government block. We rather expect the opposite; both caretaker governments and substantive minority governments will be forced to seek support for their bills on a vote by vote basis, which will lead to development of minimal connected winning coalitions. The observable consequence of this development will be the emergence of the left-right spectrum as the main dimension of parliamentary voting under caretaker governments and substantive minority governments (similar to Hix and Noury 2016) as suggested by the hypothesis H1b:

H1b: Left-right positions explain parliamentary voting patterns under caretaker governments and substantive minority governments.

3 Data and Methods

Focusing on the voting behavior in the Czech Republic is an asset for two reasons. First, the large number of votes and availability of the data from long period of time make it possible to conduct a longitudinal analysis of voting patterns in a very stable context. Existing studies often rely on a snapshot of selected cases from different countries and different years. Such case selection may cast doubts on the validity of the comparisons. In our analysis, we can directly test when and how legislative voting patterns change after a change of the type of government. Second, all votes in the Czech Republic are recorded. This minimizes the potential selection bias due to strategic roll-call vote requests which has been discussed for other parliaments (Carrubba et al. 2006; Hug 2010). However, particularly in older legislative periods, not all information about the votes is available. Yet, it is important to highlight that this is a *technical* limitation concerning relatively few votes and the absence is not caused by parliamentary rules.⁴ We can then assume these votes are missing randomly. In the next section we describe in more detail how the data was collected. Then we outline our methodological approach.

3.1 Data

This paper relies on a novel data set covering all roll-call votes from the lower house of the Czech Parliament (Chamber of Deputies) in the period from 1996 to 2014. Information on voting behavior of legislators is publicly available on the website of the parliament.⁵ We collected all individual vote decisions by legislators using web scraping techniques implemented in R (see e.g. Munzert et al. 2015).⁶

Our data set includes a total of 46,539 votes from five legislative periods. The first Chamber of Deputies is not included in the analysis, since its members

⁴In other words, all votes were originally recorded but the results were not uploaded to the webpage for whatever reason.

⁵E.g.: <http://www.psp.cz/sqw/hlasysqw?g=45031>.

⁶The data is available at request for the purpose of replication and will be made publicly available in the near future.

were elected before dissolution of Czechoslovakia and the party system was at that time still quite fluid. It was also characterized by high incidence of missing records. The data from all legislative terms included in the analysis was subdivided into smaller data sets if a change in the composition of the government occurred between the elections. This was true in all legislative periods except for the third Chamber of Deputies, when a Minority government of Zeman survived complete four years. In the empirical analysis we offer additional detailed information on the analyzed cabinets in order to put the observed results into context.

Table 2 summarizes descriptive statistics related to the roll-call data and Table 3 provides more detailed information on the type of governments. During the existence of most cabinets five parties (plus independent legislators) were represented in the Chamber of Deputies. The major factions which held seats in all terms were the Czech Social-Democratic Party (ČSSD), the conservative Civic Democratic Party (ODS) and the Communist Party of Bohemia and Moravia (KSČM). The centrist Christian-Democratic Union-Czechoslovak People's Party (KDU-ČSL) won seats in all terms except for the sixth. It managed to reenter the Parliament in 2013 (the 7th term which started in 2013 is not included in our analysis), therefore we also provide detailed information on its representatives appearing in the roll-call matrix. In the second legislative period the

Table 2 Overview of roll-call data

PM	Term	Votes	Legislators	Parties	ČSSD	ODS	KSČM	KDU-ČSL	Others
Klaus 2	2	3186	217	6	61	75	22	20	39
Tošovský	2	1395	240	7	58	71	22	19	70
Zeman	3	13977	209	5	75	65	24	22	23
Špidla	4	6689	210	5	72	60	42	22	14
Gross	4	2235	206	5	76	57	41	21	11
Paroubek	4	4689	210	5	78	58	42	21	11
Topolánek 1	5	925	203	5	74	81	26	13	9
Topolánek 2	5	5046	219	5	79	83	27	14	16
Fischer	5	2499	218	5	78	81	26	15	18
Nečas 1	6	3405	213	5	60	54	26	0	73
Nečas 2	6	2257	221	5	61	56	26	0	78
Rusnok	6	236	202	5	54	50	26	0	72

Table 3 Overview of Czech Governments (1996–2013)

PM	Party (PM)	Leg. Period	From	To	Majority (Ministers)	Majority (Investiture)	Type
Klaus 2	ODS	2	04.07.1996	02.01.1998	0	0	F. Minority
Tošovský	Ind.	2	02.01.1998	22.07.1998	0	1	Caretaker
Zeman	ČSSD	3	22.07.1998	15.07.2002	0	0	S. Minority
Špidla	ČSSD	4	15.07.2002	04.08.2004	1	1	Majority
Gross	ČSSD	4	04.08.2004	25.04.2005	1	1	Majority
Paroubek	ČSSD	4	25.04.2005	04.09.2006	1	1	Majority
Topolánek 1	ODS	5	04.09.2006	09.01.2007	0	0 ^a	S. Minority
Topolánek 2	ODS	5	09.01.2007	08.05.2009	0	0 ^b	F. Minority
Fischer	Ind.	5	08.05.2009	13.07.2010	0	1	Caretaker
Nečas 1	ODS	6	13.07.2010	17.04.2012	1	1	Majority
Nečas 2	ODS	6	17.04.2012	10.07.2013	1	1	Majority
Rusnok	Ind.	6	10.07.2013	25.11.2013	0	0 ^a	S. Minority

Note: Majority (Ministers) indicates whether parties belonging to the coalition had more than 100 (of 200) seats in the Chamber. Majority (Investiture) indicates whether a government was supported by more than 100 *MPs in the initial vote of confidence*.

^aMotion rejected. ^bThe motion was supported by exactly 100 of 200 MPs

number of parties represented in the Chamber was particularly high. Besides the established factions also a far-right Coalition for Republic-Republican Party of Czechoslovakia (SPR-RSC) and the minor conservative Civic Democratic Alliance (ODA) had parliamentary representation during the existence of the second Klaus government. Under the caretaker government of Tošovský they were joined by the liberal Freedom Union (US), a splinter of the ODS, which re-entered the parliament also during the existence of the cabinet of Zeman. In the fourth term the party was renamed into US-DEU after a merger with the Democratic Union. In 2006 it did not pass the electoral threshold. Under Topolánek 1, Topolánek 2, and Fischer the mainstream parties were joined instead by the Green Party (SZ). In the sixth legislative period the SZ left the parliament. Two new parties won the parliamentary representation instead, the conservative Tradition Responsibility Prosperity (TOP-09) and the populist Public Affairs (VV). Their emergence and their electoral success are reflected by the increasing numbers of legislators listed in Table 2 as “others”.

The highest number of votes was recorded under the cabinet of Zeman (13,977), the only cabinet of those taken into consideration which lasted the whole legislative period. The fewest votes, on the other hand, took place during the existence of the government of Rusnok (merely 236) which was created at the time of political turmoil following the resignation of Petr Nečas. In general we observe fewer parliamentary votes during the existence of caretaker governments (Tošovský and Fischer). They usually have transitional character, function for relatively short period of time and focus on running current affairs instead of engaging in active policy-making. Also the government by Rusnok was intended to be a caretaker government consisting of experts, it however did not pass the investiture vote was hence classified as a substantive minority government. The number of legislators included in the analyses is usually higher than 200 being members of the Chamber of Deputies at any given time. This is a result of party switching among elected legislators as well as of mergers and divisions of existing parties. A single legislator may appear than several times in the data set with different party affiliation.

3.2 Method

Our empirical analysis is based on the application of the Optimal Classification (OC) scaling technique (Poole 2000, 2005; Armstrong et al. 2014). OC is a standard technique for analyzing roll-call votes, but W-NOMINATE (Poole et al. 2011) and IDEAL (Clinton et al. 2004) are also often used alternative methods. We choose OC for two reasons. First, it performs better in parliaments with high levels of party discipline which is the case in the Czech Republic. Alternative techniques, particularly W-NOMINATE, can produce less reliable estimates under such conditions (Rosenthal and Voeten 2004; Armstrong et al. 2014). Secondly, OC is applicable even when only relatively few votes are scaled. This is important due to the specific research design we employ.

To analyze *how and when* legislative voting patterns are affected by a change in the type of government we use the following approach. First, all votes from a legislative period are sorted chronologically. In the second step we split these data sets into subsets each consisting of 200 votes.⁷ Then, in the third step, we

⁷We split data sets based on the number of votes (and not based on the date) to ensure that each subset of votes consists of a sufficient number of votes to be scaled. As the data is sorted by date the subsets of the data are still in chronological order.

apply OC to all the subsets of the data.⁸ The last step of the analysis is to take the average position on the first dimension for all major parties in each subset.⁹ These position estimates for the main political parties are then displayed next to each other in the chronological order for every single legislative period. Put differently, our empirical analysis is based on graphs in which the x-axis describes the data subsets (starting with the first 200 votes of a legislative period and so forth) and the y-axis displays the average OC position on the first dimension for each party. Since the data is in chronological order, we can display a change in the government type by a dashed vertical line for the subset in which the change took place.

Our approach has certain advantages but also some limitations which have to be discussed. The most important advantage of this method is that we can *directly* observe how voting patterns have developed over time within a legislative period. We can therefore analyze whether a change in voting patterns took place directly after a change in the government type or whether it happened earlier or later in the legislative period. This increases the validity of the analyses compared to, for example, estimating two different OC models for two different type of governments. With such an approach we would not know when the change in voting patterns actually took place, we would only be able to say that the voting patterns were indeed different. The second advantage of this approach is that we can test for the stability of legislative voting patterns. Since our theoretical expectations apply to the complete period of a specific government type it is important to know whether the expected patterns of legislative voting, e.g. government-opposition voting, can really be observed throughout the whole period of a government type or whether other factors also affect the observed voting patterns.

Our approach has also some limitations. Most importantly, the estimates of the party positions cannot be compared between the roll-call subsets. This means that, for example, an average OC party position of 0.7 in the first 200 votes and an average position of 0.8 in the next 200 votes does not necessarily imply that a party changed its position. This drawback, however, is unlikely to affect our results since we are interested in the relative position of parties *with respect to*

⁸We computed two-dimensional OC models, but focus only on the first dimension in this article since it is the most relevant. Abstentions were treated as missing values, but results remain unchanged when treated as ‘no’-votes.

⁹Since party unity is so high in the Czech Republic all legislators of a party have almost the same estimated ideal-point. We could have also taken the median legislator as the party position, but this does not affect our results.

each other within the roll-call subsets. Yet, in order to account for smaller changes in the estimates of party position, we use loess smoother, which simplifies the identification of general trends and abrupt changes in party positions (Jacoby 2000). We display the raw data in the appendix of this paper. Finally it should be noted that we provide no formal test of our hypothesis based on some additional statistical analysis. Instead, we identify the effect by interpreting the patterns that emerged from the spatial approach. Our approach, therefore, combines spatial modeling techniques with case knowledge about the specific governments.

4 Mapping the Policy Space

In this section we discuss the policy positions of legislators estimated using Optimal Classification. Our graphs have the advantage of showing the gradual shift of position over time, reflecting changes in the ruling coalition and government status. To improve the clarity of the figures we plot loess smoothers of the mean party position recorded at different points in time. We discuss each legislative period and cabinet separately, providing information on the specific circumstances which led to the government changes.

4.1 The Second Chamber of Deputies

The Czech party system stabilized quickly after the democratic transition (Shabad and Slomczynski 2004, p. 156). Already by mid-90s, the conservative ODS “had built up an organization that (...) included a nationwide network of local groups with well-resourced headquarters and professional regional and national structures” (Tavits 2012, p. 90). As a result, it became the most successful right-of-the-center party. After the election of charismatic Miloš Zeman as its leader in 1993 (Hloušek and Kopeček 2008, p. 528), the ČSSD emerged as the major competitor of the ODS. Since the second legislative period (1996–1998) these two parties have dominated Czech politics.

From 1993 to 1998 the ODS was a senior coalition partner in two governments of Klaus. As mentioned above, we do not include the first Czech Chamber of Deputies (1993–1996) due to its transitional character and relatively high incidence of missing votes. The first parliament elected after dissolution of Czechoslovakia was the second Czech Chamber of Deputies. The center-right-coalition which ruled the country before the election remained in power. It lost, however, the majority of votes in the Chamber. Still, as reflected by the positions of parties

in Fig. 1, the ruling coalition consisting of the ODS, the ODA and the KDU-ČSL remained a relatively coherent bloc for the entire legislative period. The opposition consisting of the ČSSD, the far-left KSČM and the far-right SPR-RSC appears at the lower end of the figure. Positions of parties remain relatively stable after the breakdown of the coalition and its replacement by the caretaker government headed by Josef Tošovský in January 1998.

The results of our position estimation approach confirm our expectations with regard to formal minority governments. The governing parties ODS, KDU-ČSL, and ODA show almost identical positions throughout the whole legislative period. The opposition parties are found at the other end of the first dimension. They do not form such a homogeneous bloc compared to the governing parties. This makes sense as the opposition consists of two radical (left- and right-wing) parties and the social-democrats (ČSSD). Strategic cooperation between these parties is unlikely leading to a less unified behavior of the opposition.

In contrast to the assumption of caretaker governments as being “non-partisan”, the government of Tošovský consisted of members of the ODA, the KDU-ČSL and former members of the ODS, who created a new party Freedom Union (US), hence we coded these factions as “government”. The creation of the government, nevertheless, relied on the support of the ČSSD in the investiture vote, which was officially part of the opposition (Kopecky and Mudde 1999, p. 416). However, the ČSSD supported the new government only in the investiture vote, but not in other roll-call votes. Due to the partisan character of the new cabinet, voting patterns remained stable and resemble those observed under the cabinet of Klaus.

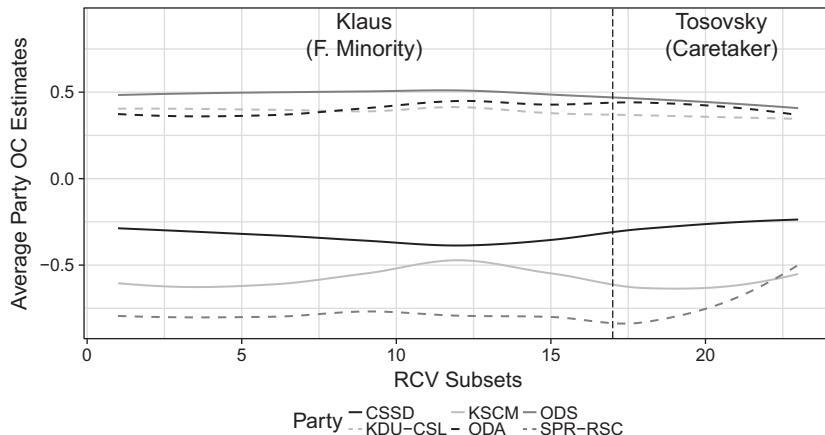


Fig. 1 Position of Czech parties in the second Chamber of Deputies (1996–1998)

4.2 The Third Chamber of Deputies

The early election in 1998 brought to power the ČSSD which became the biggest parliamentary faction. The party was, however, unable to find coalition partners and created a minority cabinet under Miloš Zeman. In spite of its minority status it proved, nevertheless, more stable than its predecessor. A reason for that was an “opposition agreement” (*opoziční smlouva*) between the ČSSD and the ODS, its principal component being “that the ODS would neither instigate nor support a no-confidence vote against the ČSSD minority government” (Plecitá-Vlachová and Stegmaier 2003, p. 773). The price for the support for the government in the investiture vote was the election of the conservative leader and former PM, Václav Klaus, to the position of the parliament’s president. The opposition agreement did not, however, guarantee support for the government in all votes, therefore the cabinet of Zeman can hence be considered as an example of *substantive* minority government.

The structure of the Czech policy space during the third legislative period conforms to what could be expected when a substantive minority government is in power. We can hence confirm findings of Hix and Noury (2016) who use the cabinet of Zeman as a case study. As displayed in Fig. 2 the dimension best explaining voting behavior of Czech legislators under the premiership of Zeman is the position of their parties on the left-right scale. The conservative ODS appears in the upper section of the plot. The liberal US is located slightly lower than the ODS and builds one cluster with the centrist KDU-ČSL. In the lower section of the graphic we find the left-of-the-center ČSSD and the far-left KSČM. Consistent with H1b, under the substantive minority cabinet of Zeman the left-right-placement of parties explains legislators’ voting behavior better than their government-opposition status.

4.3 The Fourth Chamber of Deputies

After the 2002 elections the Social-Democrats were finally able to form a majority coalition. They were joined in the government by two centrist parties, the KDU-ČSL and the US-DEU. The latter was an expanded US which merged with a minor liberal Democratic Union (DEU). The majority government survived the whole legislative period, it was, however, weakened by the resignation of the first PM, Vladimír Špidla, and the corruption scandal which led to the replacement of his successor, Stanislav Gross, by Jiří (Plecitá-Vlachová and Stegmaier 2008, p. 180).

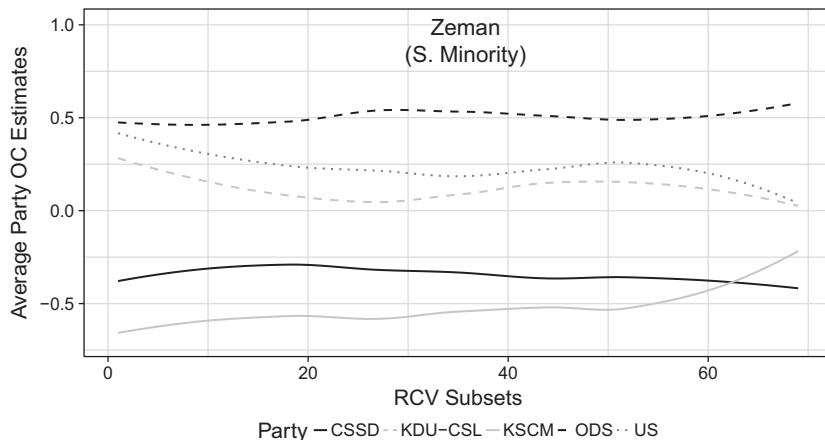


Fig. 2 Position of Czech parties in the third Chamber of Deputies (1998–2002)

The crises in the government are well reflected by party positions displayed in Fig. 3. Under Špidla we can observe the expected government-opposition divide in legislative party positions. The three parties of the ruling coalition appear close to each other in the lower section of the graphic. The conservative ODS is positioned on the opposite side of the political spectrum. Between the center-left government and the right-of-the-center opposition appears the far-left KSČM. The placement of the KSČM should not be overemphasized since it is more anti-systemic than all other factions, which followed the policy of its isolation (Hloušek and Kopeček 2008, p. 527). Such parties do not fit easily into the existing policy space (Bakker et al. 2014). Moreover, radical legislators tend to vote more sincerely (Spirling and McLean 2007), which means that they less often vote strategically and are more willing to express their “true” ideological position. This implies that legislators of the KSČM might sometimes prefer to vote with the government instead of casting the same vote as the conservative opposition ODS.¹⁰ Finally, a more detailed analysis reveals that KSČM is the only party relevant for the second OC dimension (see Appendix).

¹⁰In the appendix all plots are provided with the observed data points and without loess smoother. Here it becomes even clearer that in some roll-call vote subsets the KSČM voted with the ODS whereas and in others KSČM is placed between the government and opposition. This corroborates our impression that KSČM did not follow a clear voting pattern. With the cabinet change from Špidla to Gross, however, KSČM voting patterns shifted.

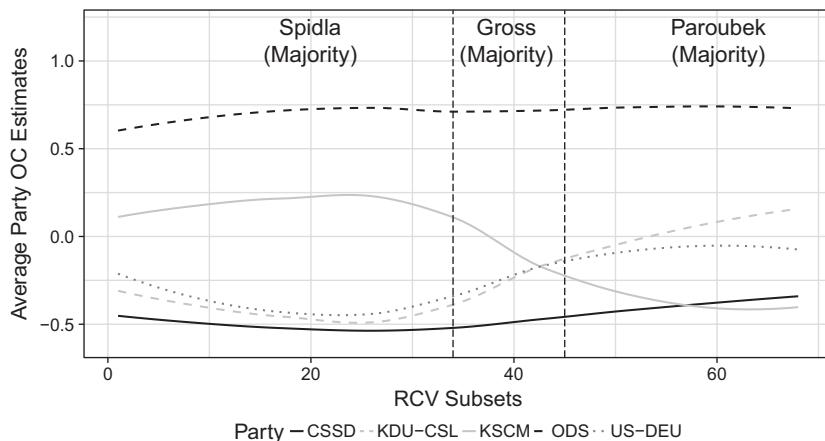


Fig. 3 Position of Czech parties in the fourth Chamber of Deputies (2002–2006)

The differences between the KSČM and all other parties on the second dimension diminished after Communists changed their voting behavior with the beginning of the Gross cabinet. Simultaneously, the minor coalition parties shifted their positions on the first dimension towards the conservative opposition. Under Paroubek this process is completed, as parties occupy the first dimension with respect to positions consistent with their left-right placement.

This observed shift in party positions is inconsistent with the theoretical expectation of government-opposition voting under majority governments. Instead, these findings have to be explained by the severe crisis of the coalition under the premiership of Gross. The crisis led to a situation which might be best described as a “quasi caretaker government” under Gross and Paroubek. The coalition partners agreed on building the cabinet but they distrusted each other, which might be best illustrated by the severe crisis under Gross “when the KDU-ČSL threatened to leave the coalition unless Gross resigned” (Plecitá-Vlachová and Stegmaier 2003, p. 180). The example of the position shift observed among parties in the fourth Czech Council of Deputies is interesting as it indicates that the effect of institutional factors may sometimes be suppressed by “external shocks” such as corruption scandals. It also displays that party positions in legislative voting are not always stable and that focusing on specific time periods leads to new insights into legislative position taking.

4.4 The Fifth Chamber of Deputies

After the parliamentary election 2006 the conservative ODS became the strongest party, but it initially did not manage to create a coalition government. The consequence was an emergence of the short-lived substantive minority government headed by the leader of the ODS, Mirek Topolánek, which existed from September 2006 to January 2007. In that period, the positions of parliamentary factions correspond roughly with their placement on the left-right-scale. As observable on the left-hand side of the Fig. 4, the ruling ODS is positioned next to the KDU-ČSL. The opposite pole consists of two left-of-the-center parties, the ČSSD and the KSČM. In the middle of the policy space appears the environmentalist SZ.

The establishment of the coalition government consisting of the ODS, the KDU-ČSL and the SZ resulted in the emergence of a structure consisting of two party blocs. The second cabinet of Topolánek was, however, formally still a minority government as it controlled exactly 50% of seats in the Chamber of Deputies. The most interesting development observed between Topolánek's first and second cabinet is the shift of the SZ away from the left-of-the-center parties and towards its new coalition partners. In spite of the formal minority status the voting behavior of legislators resembles that found under majority governments.

The second cabinet of Topolánek survived four votes of no-confidence, but lost narrowly the fifth no-confidence vote in March 2009 (Stegmaier and Vlachová 2011, p. 238). It was then replaced by the caretaker government under

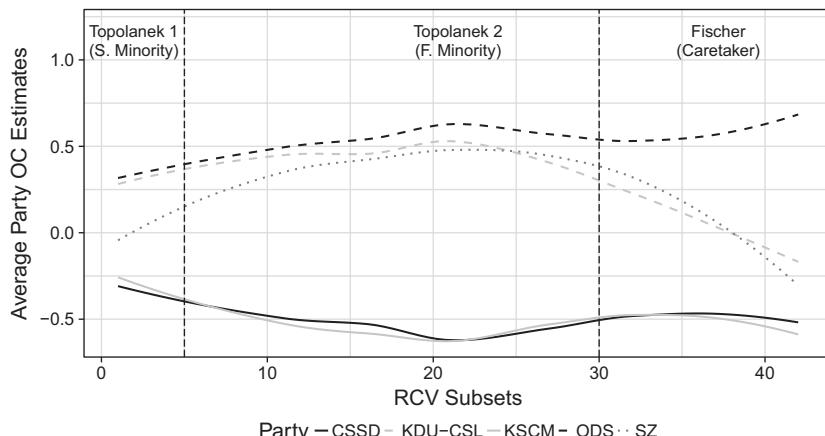


Fig. 4 Position of Czech parties in the fifth Chamber of Deputies (2006–2010)

the premiership of Jan Fischer. During that period parties shifted their positions. Their left-right placement offers once more the best explanation of legislators' voting behavior. The conservative ODS and the communist KSČM occupy now the opposite poles of the policy space. The centrist KDU-ČSL and the Greens move towards the Social-Democrats who are placed slightly closer to the center of the policy space than the Communists.

4.5 The Sixth Chamber of Deputies

The sixth Chamber of Deputies is the last Czech parliament for which we collected complete voting records. During this legislative period we may observe the effects of changes in the composition of the ruling coalition which, however, did not result in change of the government status. After the 2010 parliamentary elections the ODS formed a coalition with two new parties, the liberal-conservative TOP-09 and the populist law-and-order party VV. The leader of the ODS, Petr Nečas, became the Prime Minister. Due to the erosion of popular support for the VV, its two cabinet ministers resigned their cabinet posts and pressured the party to exit the coalition (Stegmaier and Linek 2014, p. 386). It led to a split in the party. Former members of the VV created a new grouping, Liberal-Democrats (LIDEM) (recorded in parliamentary votes as independents) which replaced the VV in the coalition with the ODS and the TOP-09. This change is reflected by the dramatic shift in the position of the remaining VV whose legislators move towards the positions of the left-of-the-center opposition (Fig. 5). Both under

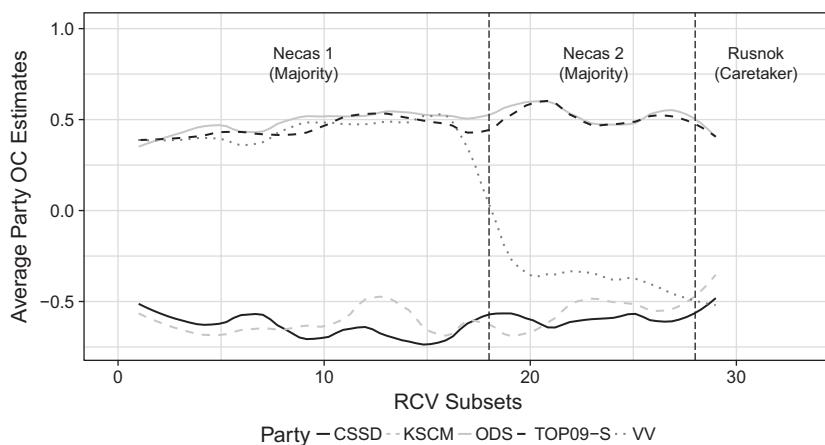


Fig. 5 Position of Czech parties in the sixth Chamber of Deputies (2010–2013)

Nečas 1 and Nečas 2 the government-opposition distinction explains legislators' voting decisions better than the left-right-placement of parties.

The distinction between the ODS and the TOP-09 on the one hand and the ČSSD, the KSČM and the VV on the other continues under the substantive minority Rusnok government. The observed continuity is unexpected and contrasts with our theoretical expectations. The three former opposition parties which supported Rusnok's motion for parliament's confidence (investiture vote) behave similarly while the former government parties still demonstrate high coherence. While thinking about this finding, we must, however, take into consideration some special features of the Rusnok cabinet. First, it failed to obtain parliament's support in the investiture vote (Stegmaier and Linek 2014, p. 386), which contributed to its extreme weakness. It is hence an example of an aborted attempt to form a caretaker government which ended up becoming a substantive minority cabinet. Second, the estimates of the policy positions of parties during the existence of the Rusnok cabinet are based on merely 236 votes. This number is substantively lower than in any other period under consideration and may result in lower reliability of estimates.

5 Theoretical Implications and Conclusion

Most of the observed positions of Czech parties correspond with our expectations about voting patterns under different government types. Under substantive minority governments and most caretaker governments the explanatory power of the left-right placement of parties surpasses the explanatory power of the variable reflecting their government-opposition status. The opposite is true in case of most majority governments and all formal minority governments. We find only two noteworthy exceptions from the aforementioned rule. The first is the majority coalition under Paroubek and the second is the substantive minority government we describe as Rusnok. In both cases important external factors, which were not related to the government status, may be responsible for the deviant results. External shocks such as distrust between coalition partners due to past corruption scandal or bitter animosity between VV and its former coalition partners modify the structure of the policy space unexpected direction (Paroubek) or result in lack of expected change (Rusnok). In case of Rusnok we must also mention the low number of scalable votes we could use to obtain our OC estimates.

Except for the above mentioned cases our results are consistent with the theoretical expectations formulated in H1a and H1b. These results imply that it is not the formal status of the government, but its effective reliance on stable parliamentary majorities. The majority should furthermore be constantly available rather than

emerge just on one occasion, e.g. at the investiture vote. Furthermore, the results have more general implications for the analysis of parliamentary voting. The extreme shifts of party positions due to government turnovers imply that it is important to estimate party positions based on the respective government type and not for complete legislative periods. In the latter case, different voting patterns have to be fit into a single model, which can increase the dimensionality and make the interpretation of party positions more challenging. However, our results suggest that voting patterns during a cabinet's term are quite stable, implying that legislative voting is mostly influenced by the relationship of a party to the government. Only the governments of Gross and Paroubek were strongly affected by a internal crisis of the coalition despite a majority government. Further, although government types can explain voting patterns very often, they do not determine them. This is highlighted by the observed exceptions to the expected voting patterns, which could be traced back to the specific political circumstances.

Our results also suggest that it makes sense to differentiate between subtypes of minority governments based on their ability to create relatively stable majorities. Only if this is not the case, can we expect voting patterns in line with left-right differences. Of course, distinguishing between formal and substantive minority governments must almost always be done based on an in-depth case analysis, taking into consideration all of the relevant information about the respective cabinet.

What's more, the analysis contributes to the discussion of a less studied type of government, the caretaker government. Our results indicate that cooperation between many factions, which is necessary for the emergence of caretaker governments, does not necessarily imply the emergence of a government-opposition divide. On the contrary, a caretaker government is rather a marriage of convenience. During its existence, parties' parliamentary voting patterns follow ideological considerations and are less strategically motivated. But just like with the analysis of minority governments, caretaker governments can take different forms. Some might be more partisan than others and this influences legislative behavior.

Finally, the limitations of our study have to be highlighted. We focus in this paper on only one country in order to increase the explanatory power by keeping the overall context as stable as possible. This may, however, raise questions about the generalizability of our results to other political contexts. Therefore, we advocate more research on longitudinal legislative voting patterns in other countries.

Appendix

See Figs. 6, 7, 8, 9, 10 and 11.

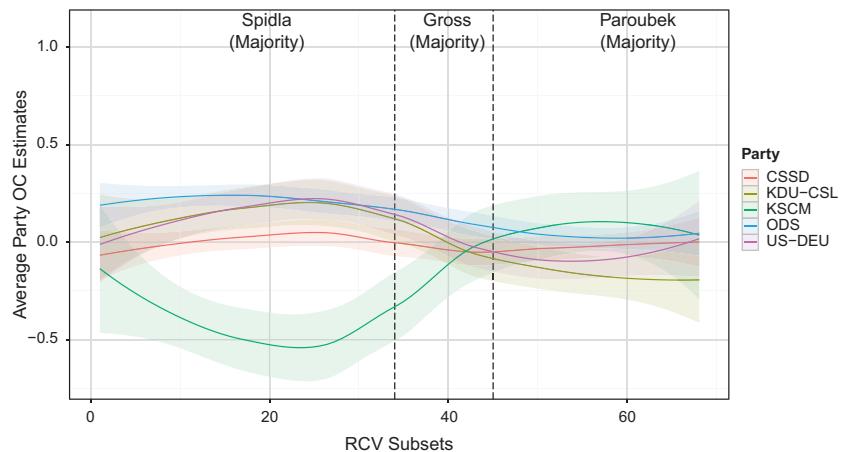


Fig. 6 Second dimension in the fourth Chamber of Deputies

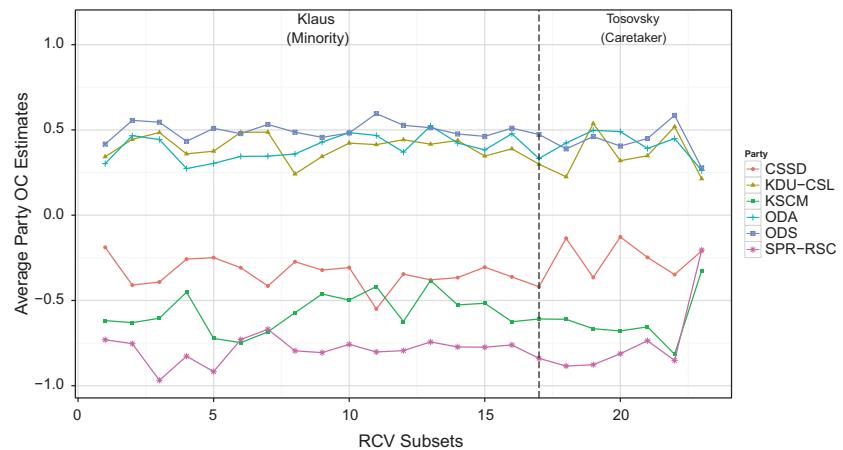


Fig. 7 Raw (no loess) positions of Czech parties in the second Chamber of Deputies

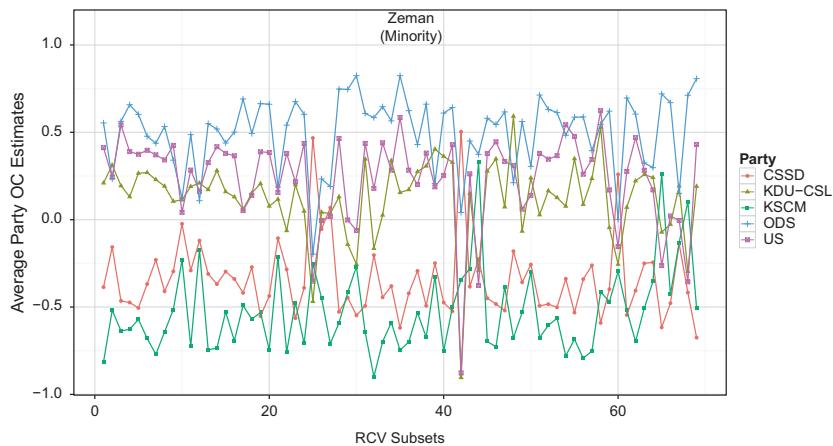


Fig. 8 Raw (no loess) positions of Czech parties in the third Chamber of Deputies

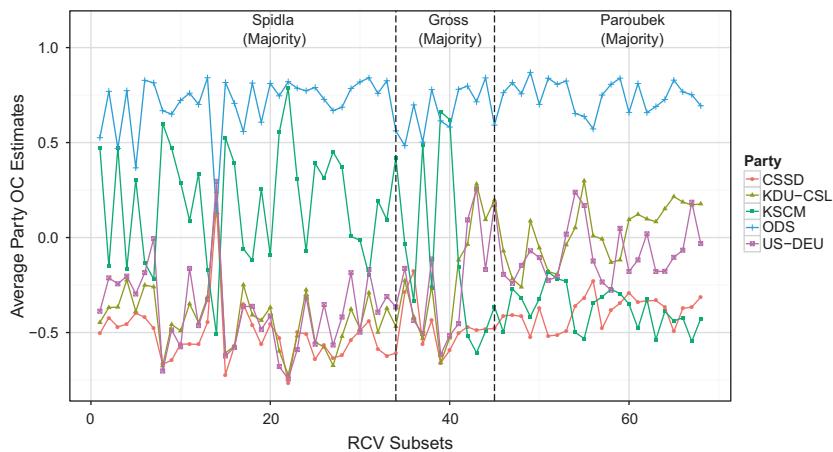


Fig. 9 Raw (no loess) positions of Czech parties in the fourth Chamber of Deputies

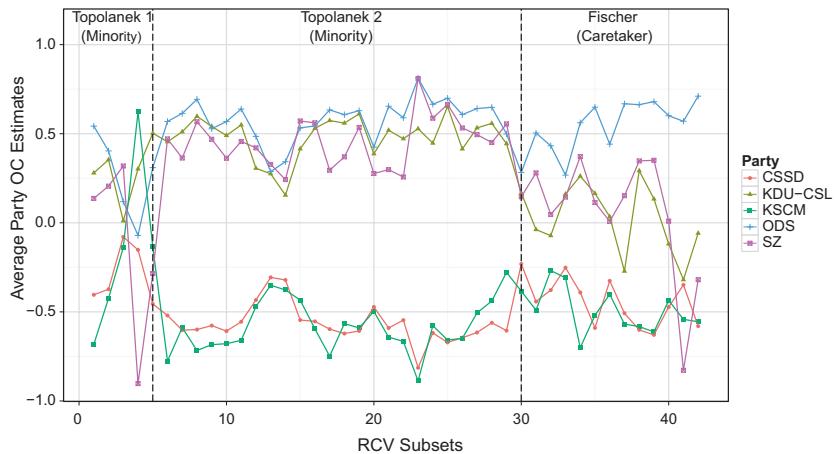


Fig. 10 Raw (no loess) positions of Czech parties in the fifth Chamber of Deputies

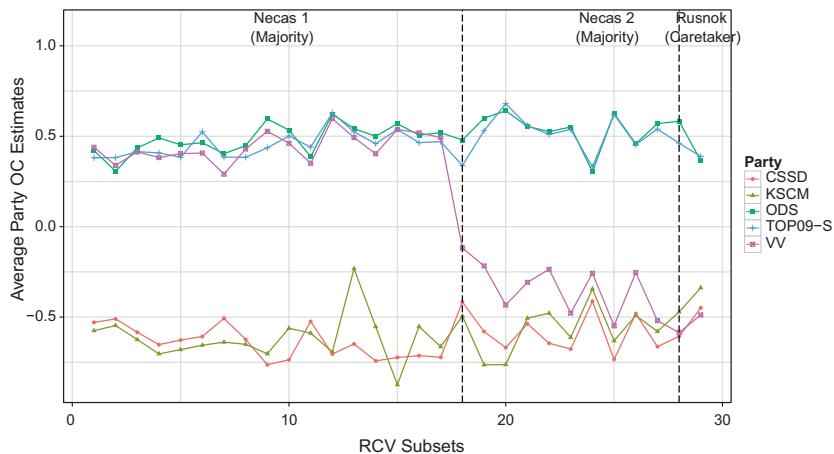


Fig. 11 Raw (no loess) positions of Czech parties in the sixth Chamber of Deputies

References

- Armstrong, David A, Ryan Bakker, Royce Carroll, Christopher Hare, Keith T Poole and Howard Rosenthal. 2014. *Analyzing spatial models of choice and judgment with R*. CRC Press.
- Axelrod, Robert. 1970. *Conflict of interest*. Markham.
- Bakker, Ryan, Seth Jolly, Jonathan Polk and Keith Poole. 2014. "The European Common Space: Extending the Use of Anchoring Vignettes." *The Journal of Politics* 76(4):1089–1101.
- Bräuninger, Thomas, Jochen Müller and Christian Stecker. 2016. "Modeling Preferences Using Roll Call Votes in Parliamentary Systems." *Political Analysis* 24(2):189–210.
- Carey, John M. 2007. "Competing Principals, Political Institutions, and Party Unity in Legislative Voting." *American Journal of Political Science* 51(1):92–107.
- Carroll, Royce and Keith Poole. 2014. "Roll call analysis and the study of legislatures." *The Oxford handbook of legislative studies* pp. 103–124.
- Carrubba, Clifford J, Matthew Gabel, Lacey Murrah, Ryan Clough, Elizabeth Montgomery and Rebecca Schambach. 2006. "Off the record: unrecorded legislative votes, selection bias and roll-call vote analysis." *British Journal of Political Science* 36(4):691–704.
- Christiansen, Flemming Juul and Helene Helboe Pedersen. 2014a. "Minority coalition governance in Denmark." *Party Politics* 20(6):940–949.
- Christiansen, Flemming Juul and Helene Helboe Pedersen. 2014b. "Minority coalition governance in Denmark." *Party Politics* 20:940–949.
- Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.
- Coman, Emanuel. 2015. "Dimensions of political conflict in West and East: An application of vote scaling to 22 European parliaments." *Party Politics*.
- Conrad, Courtenay Ryals and Sona N. Golder. 2010. "Measuring government duration and stability in Central Eastern European democracies." *European Journal of Political Research* 49(1):119–150.
- Cox, Gary. 2001. "Introduction to the Special Issue." *Political Analysis* 9(3):189–191.
- Dewan, Torun and Arthur Spriling. 2011. "Strategic Opposition and Government Cohesion in Westminster Democracies." *American Political Science Review* 105(2):337–358.
- Godbout, Jean-François and Bjørn Høyland. 2011a. "Coalition voting and minority governments in Canada." *Commonwealth & Comparative Politics* 49(4):457–485.
- Godbout, Jean-François and Bjørn Høyland. 2011b. "Legislative voting in the Canadian Parliament." *Canadian Journal of Political Science* 44(02):367–388.
- Hix, Simon, Abdul G Noury and Gérard Roland. 2007. *Democratic politics in the European Parliament*. Cambridge University Press.
- Hix, Simon and Abdul Noury. 2016. "Government-Opposition or Left-Right? The Institutional Determinants of Voting in Legislatures." *Political Science Research and Methods* 4:249–273.
- Hix, Simon, Abdul Noury and Gerard Roland. 2006. "Dimensions of politics in the European Parliament." *American Journal of Political Science* 50(2):494–520.
- Hloušek, Vít and Lubomír Kopeček. 2008. "Cleavages in the Contemporary Czech and Slovak Politics. Between Persistence and Change." *East European Politics and Societies* 22(3):518–552.
- Hloušek, Vít and Lubomír Kopeček. 2014. "Caretaker Governments in Czech Politics: What to Do about a Government Crisis." *Europe-Asia Studies* 66(8):1323–1349.

- Hug, Simon. 2010. "Selection effects in roll call votes." *British Journal of Political Science* 40(1):225–235.
- Hug, Simon. 2013. *Parliamentary Voting*. New York: Springer pp. 137–157.
- Jacoby, William G. 2000. "Loess: a nonparametric, graphical tool for depicting relationships between variables." *Electoral Studies* 19(4):577–613.
- Kalandrakis, Tasos. 2015. "A theory of minority and majority governments." *Political Science Research and Methods* 3(2):309–328.
- Kam, Christopher. 2014. "Party Discipline." *The Oxford Handbook of Legislative Studies* p. 399.
- Kam, Christopher J. 2009. *Party discipline and parliamentary politics*. Cambridge University Press.
- Kopecky, Petr and Cas Mudde. 1999. "The 1998 parliamentary and senate elections in the Czech Republic." *Electoral Studies* 18(3):415–424.
- Laver, Michael and Norman Schofield. 1992. *Multiparty Government. The Politics of Coalition in Europe. (Comparative European Politics)*. Michigan University Press.
- Louwerse, Tom, Simon Otjes, David M Willumsen and Patrick Öhberg. 2017. "Reaching across the aisle: Explaining government–opposition voting in parliament." *Party Politics* 23:746–759.
- Lyons, Pat and Tomáš Lacina. 2009. "An examination of legislative roll-call voting in the Czech Republic using spatial models." *Sociologický časopis/Czech Sociological Review* 45(6):1155–1190.
- Martin, Shane, Thomas Saalfeld and Kaare W Strøm. 2014. *The Oxford Handbook of Legislative Studies*. Oxford University Press.
- Mitchell, Paul and Benjamin Nyblade. 2008. *Government formation and cabinet type in parliamentary democracies*. Oxford UK: Oxford University Press pp. 201–236.
- Müller-Rommel, Ferdinand, Katja Fettelschoss and Philipp Harfst. 2004. "Party government in Central Eastern European democracies: A data collection (1990–2003)." *European Journal of Political Research* 43(6):869–894.
- Munzert, Simon, Christian Rubba, Peter Meißner and Dominic Nyhuis. 2015. *Automated data collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons.
- Müller, Jochen and Christian Stecker. 2014. *Zur Aussagekraft von Idealpunktschätzungen in parlamentarischen Systemen. Eine Analyse potenzieller Auswahlverzerrungen*. Wiesbaden: Springer VS pp. 85–111.
- Noury, Abdul and Elena Mielcova. 2005. "Electoral Performance and Voting Behavior in the Czech Republic." *Institute of Governmental Studies, University of California, Berkeley, Working Paper*.
- Plecitá-Vlachová, Klára and Mary Stegmaier. 2003. "The chamber of deputies election, Czech Republic 2002." *Electoral Studies* 22(4):772–778.
- Plecitá-Vlachová, Klára and Mary Stegmaier. 2008. "The parliamentary election in the Czech Republic, June 2006." *Electoral Studies* 27(1):179–184.
- Poole, Keith T. 2000. "Nonparametric unfolding of binary choice data." *Political Analysis* 8(3):211–237.
- Poole, Keith T. 2005. *Spatial models of parliamentary voting*. Cambridge University Press.
- Poole, Keith T and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York. Oxford University Press.

- Poole, Keith T, Jeffrey B Lewis, James Lo and Royce Carroll. 2011. “Scaling roll-call votes with w-nominate in R.” *Journal of Statistical Software* 42(14):1–21.
- Rosenthal, Howard and Erik Voeten. 2004. “Analyzing roll calls with perfect spatial voting: France 1946–1958.” *American Journal of Political Science* 48(3):620–632.
- Shabad, Goldie and Kazimierz M Slomczynski. 2004. “Inter-party mobility among parliamentary candidates in post-communist East Central Europe.” *Party Politics* 10(2):151–176.
- Sieberer, Ulrich. 2006. “Party unity in parliamentary democracies: A comparative analysis.” *The Journal of Legislative Studies* 12(2):150–178.
- Spirling, Arthur and Iain McLean. 2007. “UK OC OK? Interpreting optimal classification scores for the UK House of Commons.” *Political Analysis* 15(1):85–96.
- Stecker, Christian. 2011. “Bedingungsfaktoren der Fraktionsgeschlossenheit. Eine vergleichende Analyse der deutschen Länderparlamente.” *Politische Vierteljahresschrift* 52:424–447.
- Stegmaier, Mary and Klára Vlachová. 2011. “The parliamentary election in the Czech Republic, May 2010.” *Electoral Studies* 30(1):238–241.
- Stegmaier, Mary and Lukáš Linek. 2014. “The parliamentary election in the Czech Republic, October 2013.” *Electoral Studies* 35:385–388.
- Strøm, Kaare. 1984. “Minority Governments in Parliamentary Democracies The Rationality of Nonwinning Cabinet Solutions.” *Comparative Political Studies* 17(2):199–227.
- Strøm, Kaare. 1990. *Minority government and majority rule*. Cambridge University Press.
- Tavits, Margit. 2012. “Organizing for Success: Party Organizational Strength and Electoral Performance in Postcommunist Europe.” *The Journal of Politics* 74(1):83 – 97.
- Tsebelis, George. 1995. “Decision making in political systems: Veto players in presidentialism, parliamentarism, multicameralism and multipartyism.” *British journal of political science* 25(3):289–325.
- Voeten, Erik. 2000. “Clashes in the Assembly.” *International Organization* 54(02):185–215.



Zur Spezifizierung von Risiko und Unsicherheit in räumlichen Modellen

Guido Tiemann

Zusammenfassung

Beiträge zur räumlichen Theorie des Wählens gehen zumeist von idealisierten Bedingungen aus: Wähler sind vollständig informiert und entscheiden sich strikt rational, Parteien beziehen klare und eindeutig identifizierbare Positionen im politischen Wettbewerbsraum, Wählerinstellungen zum Umgang mit Risiko und Unsicherheit sind *a priori* in theoretischen und statistischen Modellen fixiert. Dieser Beitrag hinterfragt diese Grundannahmen der „Neo-Downsianischen“ Modelltradition. Er bestimmt empirisch, wie Wähler räumliche Distanzen in Nutzenfunktionen übersetzen und wie sie dabei mit Risiko und Unsicherheit umgehen. Ein wesentlicher Aspekt betrifft dabei die Angemessenheit von konkaven oder konvexen Nutzenfunktionen, also die Frage, ob theoretische und/ oder statistische Modelle Verlustfunktionen mit quadratischen oder mit linearen Metriken spezifizieren sollten. Die empirische Analyse verwendet das umfangreiche Datenmaterial des Wahlforschungsprojekts „The Comparative Study of Electoral Systems“ (CSES). Vergleichende Analysen des Wahlverhaltens zeigen dabei eindeutig, dass Wähler über neunzig heterogene Wahlkontakte hinweg wesentlich weniger risikoavers sind als von der großen Mehrheit theoretischer und empirischer Beiträge unterstellt wird. Stattdessen zeigen dieser Beitrag, dass moderne Wähler sich im Wesentlichen risikoneutral verhalten.

G. Tiemann (✉)

Forschungsgruppe European Governance and Public Finance, Wien, Österreich
E-Mail: tiemann@ihs.ac.at

1 Einleitung

Wissenschaftliche Modelle sind Abstraktionen. Sie sollen nicht wahr, sondern beim Erreichen von deskriptiven oder kausalen Schlussfolgerungen nützlich sein, und sie gründen regelmäßig auf einer Reihe notwendiger, teils restriktiver Annahmen. Die räumliche Theorie des Wählens unterstellt zum Beispiel, dass Wähler und Parteien fixe und eindeutig identifizierbare Positionen in n -dimensionalen ideologischen und/oder politischen Räumen einnehmen und dass die Relation von Wähler- und Parteidatenpositionen das Wahlverhalten mindestens mitbestimmt.

In diesem Beitrag greife ich zwei dieser Grundannahmen auf: Die erste betrifft die Informiertheit der Wähler und, somit, die Sicherheit ihrer Entscheidung. Bei nahe alle Beiträge des Feldes unterstellen umstandslos, dass Wähler vorgegebene Dimensionen eines politischen Raums verstehen, auf jeder dieser Dimensionen bestimmte persönliche Idealpunkte einnehmen und dass sie die Positionen der wesentlichen Kandidaten und/oder Parteidatenalternativen auf denselben Dimensionen genau kennen und exakt vermessen können. Eine ganze Reihe von Gründen spricht jedoch gegen diese idealtypischen Annahmen: Wähler sind unterschiedlich stark an politischen Fragen interessiert, und sie verfügen über unterschiedliche Ressourcen bei der Verarbeitung politischer Informationen. Politische Parteien haben oft ein strategisches Interesse daran, unklare (oder gar keine) Positionen zu bestimmten Sachfragen zu beziehen, sie sind manchmal intern gespalten und richten unterschiedliche programmatische Botschaften an unterschiedliche Segmente der Wählerschaft (vgl. Alvarez 1998; Bartels 1986; Brady und Ansolabehere 1989; Lupia 2016; Shepsle 1972; Tomz und Van Houweling 2009).

Die zweite Annahme betrifft die Wahrnehmung von geometrischer Nähe und Distanz in politischen Räumen und die Übersetzung dieser Maße in die jeweiligen Nutzenfunktionen einzelner Wähler. Die Auswahl einer geeigneten Distanzmetrik ist dabei kein triviales oder bloß technisches Detail, sondern sie hat weitreichende Konsequenzen für das theoretische Verständnis und die empirische Modellierung des Wahlverhaltens. In der Literatur sind vor allem Spezifikationen verbreitet, die à priori entweder quadratische und lineare Metriken unterstellen:

- Quadratische Verlustfunktionen setzen risikoaverse Wähler voraus. Sie sind mathematisch recht einfach handhabbar und bestimmen seit Langem den Status quo der „Neo-Downsianischen“ Modelltradition (vgl. anstatt vieler weiterer Beiträge, Adams et al. 2005; Davis et al. 1970; Enelow and Hinich 1984; Feddersen 1992; Merrill und Grofman 1999; Schofield and Sened 2006).

- Lineare Verlustfunktionen unterstellen dagegen das risikoneutrale Verhalten der Wähler. Der ursprüngliche Beitrag von Downs (1957) benutzt lineare Distanzen, und sie werden weiterhin von einer sichtbaren Minderheit der Beiträge zur räumlichen Theorie des Wählens aus theoretischen oder aus empirischen Erwägungen bevorzugt (vgl. zum Beispiel Degan und Merlo 2009; Eguia 2012; Grynaviski und Corrigan 2006; Humphreys und Laver 2010; Kramer 1977; Patty et al. 2009; Wittman 1973, 1977).
- Neben diesen einfachen Basisoptionen haben theoretische und empirische Beiträge eine Reihe komplexerer Alternativen diskutiert und verwendet, etwa exponentielle (Carroll et al. 2013; Poole und Rosenthal 1985) oder nicht-euklidische Verlustfunktionen (McKelvey 1976; Plott 1967). Schließlich verlassen einige weitere Modelle vollständig die Logik von Nähe und Distanz in euklidischen Räumen: Das Modell der Richtungswahl begreift politische Sachfragen als binär. Es begründet den eher affektiven und symbolischen Nutzen einer Parteialternative durch die richtungspolitische Übereinstimmung und die Intensität, mit der sie von Wählern und Parteien vertreten wird (Rabinowitz und Macdonald 1989).

Das Interesse an diesen Fragestellungen ist gleichermaßen aus methodischer wie aus inhaltlicher Perspektive wohlbeg理ndet. Die beiden oben dargestellten Grundannahmen bestimmen (1) wie und in welchem Maße Unsicherheit über die Position bestimmter Parteien in ideologischen oder politischen Räumen entsteht und (2) auf welche Weise einzelne Wähler mit dem jeweiligen Maß an Unsicherheit umgehen. Theoretische wie statistische Modelle müssen diese Prämissen adäquat formulieren, sodass substanziale Einsichten in den politischen Prozess, hier in die Determinanten des Wahlverhaltens, überhaupt erst möglich werden. Die hier nur kursorisch vorgestellten Modellspezifikationen gründen auf ganz unterschiedlichen Konzepten zum Umgang mit Risiko in Entscheidungssituationen. Diese Risikoorientierungen bestimmen wiederum wesentlich, wie Wähler mit begrenzten oder unsicheren Informationen umgehen, wie sie auf unbestimmte oder unklare Parteipositionen reagieren und wie sie Parteialternativen bewerten und letztlich auswählen.

Freilich können Wahlforscher nicht einfach à priori und axiomatisch unterstellen, welche konkreten Überlegungen Wähler anleiten, wenn sie räumliche Distanzen wahrnehmen und in bestimmte Nutzenfunktionen übersetzen. Und diese Frage kann auch nicht (allein) am Maßstab von mathematischer Schlichkeit und praktischer Verwendbarkeit entschieden werden. Vielmehr ist die Auswahl einer geeigneten Distanzmetrik eine letztlich empirische Frage, die durch die Analyse geeigneter Wahlstudien beantwortet und gerechtfertigt werden muss.

Beim gegenwärtigen Stand der Forschung sind nicht viele Autoren diesen Weg gegangen: Am Material amerikanischer Präsidentenwahlen haben zum Beispiel Berinsky und Lewis (2007), Grynавiski und Corrigan (2006) und Ye et al. (2011) gezeigt, dass lineare Distanzen den Datenstrukturen besser entsprechen und eine höhere Güte der Schätzung ermöglicht haben als quadratische Nutzenterme. Bei ähnlich angelegten Analysen deutscher Bundestagswahlen haben Thurner (2000) und Shikano und Behnke (2009) diese Befunde im Wesentlichen bestätigt. Schließlich hat Singh (2014) die bislang einzige ländervergleichende Studie vorgelegt, die am Material der ersten beiden Wellen des CSES-Projekts zeigt, dass lineare Nutzenterme, wieder gemessen an der Güte der Schätzung, bessere Erklärungen bieten für die Wahlbeteiligung und die Wiederwahl von Amtsinhabern.

In diesem Beitrag möchte ich über die bisherigen Fallstudien hinausgehen und gleichermaßen theoretische Überlegungen anstellen und empirische Befunde vollständig spezifizierter räumlicher Modelle präsentieren. Diese systematisch-vergleichende Analyse greift auf umfangreiche Daten der ersten drei Wellen des Projekts „The Comparative Study of Electoral Systems“ zurück (CSES; <http://cses.org/>). Substanziell konzentriere ich mich auf Wahlen zu nationalen Parlamenten, weil andere Urnengänge, etwa Präsidentenwahlen, oft eine ganz unterschiedliche Salienz aufweisen, meist viel deutlicher personalisiert und deshalb kaum sinnvoll vergleichbar sind. Die empirischen Befunde unterstreichen, über eine weite Bandbreite unterschiedlicher politischer Kulturen, Parteiensysteme und institutioneller Systeme hinweg, dass einzelne Wähler und ganze Elektorate sehr viel weniger risikoavers sind als bislang theoretisch angenommen und in den meisten Modellen empirisch spezifiziert wird. Über neunzig unterschiedliche Parlamentswahlen hinweg identifizieren die Ergebnisse nicht einen Fall, bei dem die Wähler derart risikoscheu sind wie es das „Neo-Downsianische“ Standardmodell quadratischer Verlustfunktionen unterstellt. Vielmehr tendieren die meisten Elektorate hin zu einem risikoneutralen Verhalten.

Um diese Argumentation vorzustellen und empirisch abzusichern, gehe ich in drei wesentlichen Schritten vor: Zunächst wiederhole ich einige konzeptionelle Schlüsselbegriffe und theoretische Grundlagen der räumlichen Theorie des Wählens. Dabei betone ich besonders die Vermessung räumlicher Nähe und Distanz und die theoretische Eigenschaften der Nutzenfunktionen (Abschn. 2). Im Anschluss stelle ich das empirische Datenmaterial vor, definiere und operationalisiere Schlüsselvariablen des räumlichen Modells und einige nicht-räumliche Kontrollvariablen. Die folgende empirische Analyse verwendet ausschließlich Daten der Umfragemodule, die in den ersten drei Wellen des Projekts „The Comparative Study of Electoral Systems“ abgefragt und erhoben wurden

(Abschn. 3). Im Anschluss stelle ich im Detail die empirischen Modelle der Wahlentscheidung vor und präsentiere vergleichende Befunde zur Rolle von Risikoorientierungen bei der Wahlentscheidung (Abschn. 4). Das Fazit fasst wesentliche Resultate zusammen, diskutiert ihre Implikationen für die theoretische und empirische Wahlforschung und skizziert Grundrisse eines zukünftigen Forschungsprogramms.

2 Grundannahmen des räumlichen Modells

Dieser Abschnitt beginnt mit einer kurzen Wiederholung einiger Grundlinien und zentraler Bausteine der räumlichen Theorie des Wählens. Wie bereits ausgeführt unterstelle ich, dass Wähler- und Parteipositionen in einem politischen Wettbewerbsraum sinnvoll verortet, dargestellt und aufeinander bezogen werden können. Um die theoretischen und statistischen Modelle nicht zu komplex werden zu lassen, beziehe ich mich jedoch in diesem Beitrag allein auf *eindimensionale* politische Räume.

Deterministische und probabilistische Versionen der räumlichen Theorie stimmen darin überein, dass ein Wähler v sich für denjenigen Kandidaten oder diejenige Parteialternative p entscheidet, von dem/der er sich den höchsten Nutzen $u(v, p)$ verspricht. Die rationale Nutzenkalkulation der Wähler enthält deshalb eine Verlustfunktion. Sie definiert, wie der „Nutzen“, den jeder Wähler i jeder Partei j zuschreibt, mit ihrer räumlichen Distanz im Koordinatensystem sinkt:

$$u(v, p) = f(v, p) = -\|v - p\| = -\|v_i - p_{i,j}\|. \quad (1)$$

Die Variablen v_i und $p_{i,j}$ benennen hier individual- und alternativenspezifischen Wähler- und Parteipositionen innerhalb des eindimensionalen politischen Raums. Zum Beispiel bezeichnet v_i den Idealpunkt von Wähler i , während $p_{i,j}$ die idiosynkratisch wahrgenommene Verortung von Partei j durch Wähler i angibt. $\|\bullet\|$ ist eine beliebige Distanzfunktion, die bestimmt, wie Entfernung im politischen Raum in wähler- und partiespezifische Nutzenterme übertragen werden. Obgleich lineare und, mehr noch, quadrierte Distanzen am häufigsten zur Bestimmung von Entfernung in politischen Wettbewerbsräumen verwendet werden, ist ein binärer Vergleich dieser Alternativen dennoch unnötig restriktiv. Vielmehr können lineare ($\beta = 1$) und quadrierte Abstandsmaße ($\beta = 2$) als Sonderfälle einer allgemeineren Spezifikation $L_\beta(v, p)$ verstanden werden (vgl. zum weiterführenden Konzept der Minkowski-Distanz Eguia 2013; Humphreys und Laver 2010; Laver und Hunt 1992; Thurner 2000; Ye et al. 2011):

$$u(v, p) = -L_\beta(v, p) = -|v_i - p_{i,j}|^\beta \quad (2)$$

Ich unterstelle zunächst, dass die Wahlentscheidung und die Parameter v_i und $p_{i,j}$ fixierte und gemessene Daten sind. Der Formparameter β wird dagegen durch ein statistisches Modell geschätzt; β gibt die Form der Nutzenfunktion an und bestimmt damit die Risikoorientierung der Wähler. Die Festlegung einer konkreten Distanzmetrik geht weit über den bloß technischen oder statistischen Teil des Arguments hinaus. Inhaltlich bestimmt die Form der jeweiligen Nutzenfunktion wie ein Wähler mit Risiko und Unsicherheit umgeht. Anstatt das Risikoprofil aller (*sic!*) Wähler vorab als quadratisch ($\beta = 2$) oder linear ($\beta = 1$) festzulegen, bestimme ich in diesem Beitrag den Koeffizienten empirisch: Schätzer von $\beta > 1$ zeigen generell risikoaverse Elektorate an, $\beta \sim 1$ identifiziert risikoneutrale, und $\beta < 1$ bezeichnet risikoaffine Elektorate.

Beim eindimensionalen Parteienwettbewerb bleibt die Rangfolge der Alternativen unabhängig von der gewählten Distanzmatrix erhalten: Ordinale Nutzenfunktionen, die Reihung der Präferenzen und vorhergesagtes oder „korrektes“ Wahlverhalten werden durch die Spezifikation linearer oder quadratischer Verlustfunktionen nicht verändert. Die Auswahl der Distanzmetrik modifiziert jedoch kardinale Nutzenfunktionen: Bei linearen Verlustfunktionen fällt der Nutzen einer Partei j für Wähler i gleichmäßig mit der zunehmenden Distanz von Wähler und Parteialternative $|v_i - p_{i,j}|$. Bei quadratischen Verlustfunktionen fällt dagegen der Nutzen immer stärker ab, wenn sich das Differential von Wähler- und Parteiposition $|v_i - p_{i,j}|^2$ weiter erhöht. Lineare Distanzfunktionen reagieren deshalb vergleichsweise empfindlich auf kleinere Differenzen in der Nähe des Idealpunktes der Wähler, quadratische Verlustfunktionen fallen dagegen schroffer ab, wenn Wähler- und Parteipositionen ohnehin weit voneinander entfernt sind.

Abb. 1 charakterisiert und vergleicht risikoaverse (=konkave) und risikoaffine (=konvexe) Nutzenfunktionen in eindimensionalen politischen Wettbewerbträumen. Diese exemplarische Darstellung wirft zusätzliches Licht auf die theoretischen Implikationen unterschiedlicher Verlustfunktionen und illustriert darüber hinaus die Bedeutung von Unsicherheit über die ideologischen oder politischen Positionen der Parteien. Innerhalb eines stilisierten, eindimensionalen politischen Raums konkurrieren zwei Parteien um Wählerstimmen: p_j benennt eine präzise Partei, die durch einen Punktschätzer abgebildet wird; $p_{j'}$ benennt eine unsichere Plattform, die nicht mit Bestimmtheit wahrgenommen, sondern durch eine Wahrscheinlichkeitsverteilung unscharf charakterisiert wird:

- Das linke Teilbild trägt mit einer durchgezogenen Linie die konventionelle Annahme risikoaverser Wähler und konkaver Nutzenfunktionen ab. Bei risikoaversen Nutzenfunktionen steigt der marginale Nutzenverlust mit zunehmender räumlicher Distanz stetig an, und bei gleicher Positionierung

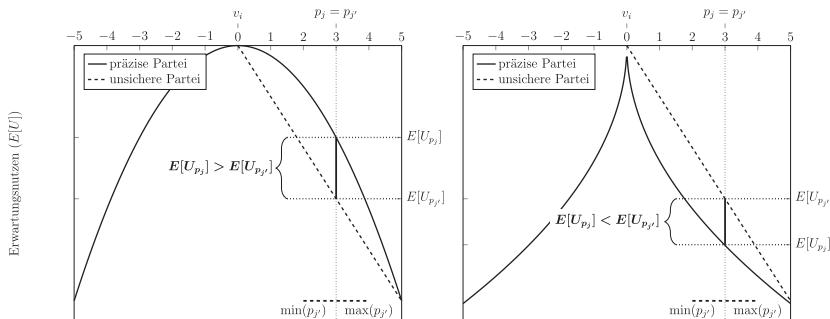


Abb. 1 Die Form der Nutzenfunktion und Risikoeinstellungen der Wähler

beider Parteien ($p_j = p_{j'}$) trägt die präzise Partei dem Wähler stets einen höheren erwarteten Nutzen ein als die unsichere Partei ($E[U_{p_j}] > E[U_{p_{j'}}]$).

- Das rechte Teilbild illustriert die Konsequenzen von Unsicherheit bei den (mindestens in Teilbereichen) konvexen Nutzenfunktionen risikoaffiner Wähler. Bei risikoaffinen Wählern nimmt der marginale Nutzenverlust hin zu den Rändern der Funktion stetig ab, und bei einer identischen (durchschnittlichen) Aufstellung beider Parteien ($p_j = p_{j'}$) erreicht die unsichere Plattform einen höheren erwarteten Nutzen als die präzise ($E[U_{p_{j'}}] > E[U_{p_j}]$).

Generell haben risikoaverse Wähler konkave Nutzenfunktionen, und ein Wähler gilt immer dann als risikoavers wenn er den erwarteten Nutzen einer präzisen Parteialternative einer fairen Lotterie mit demselben Erwartungsnutzen vorzieht. Im Gegensatz haben risikoaffine Wähler mindestens teilweise konvexe Nutzenfunktionen und bevorzugen eine unsichere Lotterie gegenüber der fixen Auszahlung ihres Erwartungsnutzens.

3 Daten und Variablen

Dieser Beitrag schätzt angemessene Distanzmetriken empirisch und bestimmt die Form der Nutzenfunktionen in räumlichen Modellen der Wahlentscheidung in einer systematisch-vergleichenden Perspektive. Die vergleichende Wahlforschung verwendet für derartige Fragestellungen meist die parallel konstruierten Umfrage-module des Projekts „The Comparative Study of Electoral Systems“ (CSES; <http://www.cses.org/>). Der umfangreiche Datenbestand des CSES-Projekts hat sich als die

Datenquelle für systematische, ländervergleichende Analysen des Wahlverhaltens etabliert. Diese Untersuchung des Wahlverhaltens unter Risiko und Unsicherheit verwendet eine Kumulation der ersten drei Wellen des CSES-Projekts. Diese Datenbasis umfasst 129 einzelne Umfragemodule zu nationalen Wahlen in mehr als vierzig Vergleichsstaaten. Konkret benutze ich den „CSES Module 1–3 Harmonized Trend File“, der am Wissenschaftszentrum Berlin von Heiko Giebler, Josephine Lichteblau, Antonia May, Reinhold Melcher, Aiko Wagner, und Bernhard Weßels zusammengestellt wurde. Daten und detaillierte Dokumentation sind über die CSES-Website verfügbar (<http://cses.org/datacenter/trendfile/trendfile.htm>; vgl. auch Anhang 6).

Die CSES-Module der unterschiedlichen Wellen enthalten alle Fragen und alle empirischen Informationen, die für die Schätzung einfacher räumlicher Modelle erforderlich sind. Die Questionnaires erheben Daten über die Wahlentscheidung, über die Idealpunkte der Wähler auf der Links-Rechts-Dimension ($v_i \in [0, 10]$) und über idiosynkratische Platzierungen der Parteialternativen auf identisch konstruierten Elfpunkteskalen ($p_{i,j} \in [0, 10]$).

In diesem Beitrag verwende ich zur Messung von Parteipositionen durchgehend idiosynkratische Platzierungen, weil jeder einzelne Wähler die Bewertung von Parteialternativen und die Auswahl unter diesen Alternativen letztlich nur von denjenigen Informationen abhängig machen kann, die ihm unmittelbar selbst vorliegen, ganz gleich ob sie nun, nach welchem Maßstab auch immer, „richtig“ oder „falsch“ sind. Mittelwerte dieser Platzierungen, die Urteile politikwissenschaftlicher Experten oder die richtungspolitische Verortung von Manifestos werden dagegen durch Informationen begründet, die einzelne Wähler regelmäßig nicht haben oder auch gar nicht haben können, und diese Maße eignen sich deshalb nur wenig zur Erklärung der individuellen Wahlentscheidung. Mit ganz ähnlicher Begründung verzichte ich darauf, Parteiplatzierungen einzelner Wähler durch weitere Skalierungsverfahren zu „korrigieren“ (vgl. etwa die Techniken bei Aldrich und McKelvey 1977; Hare et al. 2015; Lo et al. 2014) oder potenzielle Projektionseffekte zu korrigieren (vgl. Grand und Tiemann 2013; Krosnick 2002).

Neben diesen Schlüsselvariablen des räumlichen Modells erhebt das CSES-Projekt eine Reihe zusätzlicher, nicht-räumlicher Prädiktoren der Wahlentscheidung: die Art und Stärke einer langfristigen Parteiidentifikation, die Zufriedenheit mit dem jeweiligen demokratischen System, eine Batterie von drei Items, die politisches Wissen der Befragten erheben, und demografische Kontrollvariablen wie Alter, Geschlecht und Bildungsstand.

Der kumulierte Datenbestand des CSES-Trendfiles deckt den Zeitraum von 1996 bis 2011 ab. Ich verwende ausschließlich Wahlen zu nationalen Parlamenten (zu den Unterhäusern bei bikameralen Systemen), sodass die empirische Datenanalyse auf neunzig einzelnen Wahlsegmenten gründet und insgesamt 44

unterschiedliche, heterogene Vergleichsstäaten einschließt. Die Auswahl eines derart heterogenen Vergleichssets ermöglicht es, entlang der Linien eines „most different systems design“ zu prüfen, ob die Resultate einzelner Umfragesegmente deutlich durch jeweilige Kontexteffekte geformt sind oder ob die Analyse generelle Befunde erbringt, die über heterogene Kontextfaktoren hinweg „halten“ und verallgemeinerbar sind (vgl. bereits Przeworski und Teune 1970; Tiemann 2003).

Mit Blick auf die einzelnen Survey-Segmente musste ich einige sehr kleine Nachwahlbefragungen ausschließen, andere Segmente des CSES-Trendfiles konnte ich aufgrund von inkonsistenten oder fehlerhaften Daten nicht sinnvoll für die Analyse aufbereiten und verwenden. Zudem habe ich einige Umfragesegmente, die keine Angaben zur Parteiidentifikation abgefragt haben, von der empirischen Analyse ausgeschlossen. (Die für die vergleichende Analyse ausgewählten Segmente des CSES-Trendfiles sind in Anhang 6 aufgelistet und dokumentiert.).

Schließlich habe ich einige individuelle Beobachtungen nicht berücksichtigen können: Das betrifft zuerst Nichtwähler und die Unterstützer kleinerer Parteien, für die in den CSES-Questionnaires keine alternativenspezifischen Parteidpositionen erhoben werden. Weiterhin nicht in die Analyse aufgenommen werden diejenigen Befragten, die keinen Idealpunkt auf der Links-Rechts-Dimension angegeben oder eine Mehrzahl der möglichen Parteialternativen, also mehr als fünfzig Prozent der Wettbewerber bei einer Parlamentswahl, nicht im politischen Wettbewerbsraum platziert haben. Umfassender spezifizierte, gesättigte Modelle, die auch Effekte nicht-räumlicher Variablen einschließen, konnten zudem Befragte nicht berücksichtigen, die keine Angaben zur Demokratiezufriedenheit, Alter, Geschlecht oder Bildungsstand gemacht haben.

4 Die Form der Nutzenfunktion

In diesem Abschnitt übersetze ich die Forschungsfrage und die theoretischen Modelle in ein stringentes Forschungsdesign. Ich stelle zunächst den Ansatz zur empirischen Datenanalyse vor, diskutiere und rechtfertige Grundentscheidungen bei der Auswahl und Spezifizierung des statistischen Modells und beschreibe die Operationalisierung und Messung wesentlicher Modellvariablen. Im folgenden Abschnitt stelle ich die empirischen Befunde vor und zeige, dass die über die unterschiedlichen, heterogenen Kontexte des kumulierten CSES-Datensatzes hinweg Wähler eher risikoneutrale als risikoaverse Positionen einnehmen, wenn sie Parteialternativen bewerten und ihre Stimme abgeben. Sodann erweitere ich sukzessive das Modell und berücksichtige die Effekte nicht-räumlicher Kontrollvariablen und eines empirisch bestimmten Unsicherheitsindikators.

4.1 Von theoretischen zu statistischen Modellen

Das statistische Modell bildet die Grundlinien des theoretischen Arguments genau ab: Wahlforschung sollte sich mit der Wahlentscheidung beschäftigen, und die Mehrheit an einschlägigen Beiträgen versucht zu erklären, warum einer (und nur einer) der Wahlvorschläge ausgewählt wird. Ein alternativer Ansatz versteht Parteiskalometer oder Angaben zur „Propensity to Vote“ (PtV) als direkte Messung des kardinalen Nutzens (vgl. van der Eijk et al. 2006). Diese Perspektive ist somit nicht nur voraussetzungsvoller, sondern sie gründet auch auf recht unklar definiertem Material. Skalometer- oder PtV-Skalen innerhalb derselben Befragung unterscheiden sich oft deutlich, und ihre konkrete Formulierung greift anstatt kurzfristiger oder strategischer Wahl motive oft längerfristige Einschätzungen auf. Die Analyse von Skalometerdaten findet meist mit linearen Modellen statt, und diese Ergebnisse richten sich nicht direkt auf die Wahlentscheidung, sondern sie sind kontaminiert durch die Bewertungen verschiedener Parteialternativen, die in der Präferenzordnung der Wähler weit unten stehen.

Ich analysiere in diesem Beitrag deshalb nicht allgemein das Gesamtspektrum aller Parteibewertungen und -präferenzen, sondern konzentriere mich spezifisch auf die jeweilige Entscheidung eines Wählers. Deshalb verwende ich Verfahren zur Modellierung diskreter Entscheidungen, um die räumlichen und nicht-räumlichen Determinanten des Wahlverhaltens zu bestimmen, die Form der Nutzenfunktionen zu schätzen und die Unsicherheit der Befunde zu erschließen (eine allgemeine Einführung bietet Train 2009). Bei konventionellen „Conditional Logit“-Modellen der Wahlentscheidung gilt für die Wahrscheinlichkeit, dass Wähler i die Parteialternative j auswählt:

$$Pr(v_i = j) = \frac{\exp[u(v, p)]}{\sum_{j=1}^J \exp[u(v, p)]} \quad (3)$$

Dieses statistische Zufallsnutzenmodell wird durch räumliche und nicht-räumliche Prädiktoren der Wahlentscheidung operationalisiert und durch α , β und λ und θ parametrisiert. Deshalb ergibt sich der implizite Gesamtnutzen von Partei j für Wähler i mit:

$$u(v, p) = -\alpha |v_i - p_{i,j}|^\beta + \lambda_j c_j + \theta_{i,j} D_i + \epsilon_{i,j} \quad (4)$$

In einem einfachen Grundmodell ist der implizite Gesamtnutzen die Summe eines räumlichen Anteils, der auf den geometrischen Distanzen von Wählern und

Parteien im politischen Raum gründet ($|v_i - p_{i,j}|$), und einer parteispezifischen Konstante c_j , die nicht explizit modellierte und über alle Wähler(gruppen) hinweg konstante Eigenschaften einer Partei j abbildet. Im folgenden Schritt spezifizierte ich besser gesättigte Modelle und füge eine Matrix an nicht-räumlichen Einflüssen auf das Wahlverhalten ein. \mathbf{D}_i enthält als Kontrollvariablen Angaben zur Parteiidentifikation, Demokratiezufriedenheit, Indikatoren, die den Grad an politischer Information einzelner Wähler aufgreifen, das Bildungsniveau, Alter und Geschlecht jedes Wählers.

Auf der Grundlage dieser Spezifikation und der empirischen Informationen werden die Parameter des vereinigten Modells geschätzt: $\hat{\alpha}$ gibt die Salienz der räumlichen Modellkomponente an und sollte somit stets numerisch positiv und substanzial bedeutsam sein. β modelliert die Form der Nutzenfunktionen und zeigt deshalb die Risikoorientierung des Elektorats an. An dieser Stelle ist es wesentlich, daran zu erinnern, dass beide empirische geschätzte Parameter, $\hat{\alpha}$ und $\hat{\beta}$, die Salienz und die Formparameter auf der Ebene des gesamten Elektorats und nicht auf der Ebene einzelner Wähler modellieren. Mit diesem Vorbehalt identifiziert $\hat{\beta} > 1$ risikoaverse Einstellungen, $\hat{\beta} < 1$ verweist auf risikoaffine Einstellungen, und $\hat{\beta} \sim 1$ kennzeichnet vorherrschende Risikoindifferenz.

Der gleichfalls empirische bestimmte Parametervektor λ fängt unmodellierte Eigenschaften der konkurrierenden Parteien durch alternativenspezifische Konstanten auf, und der Parametervektor θ modelliert nicht-räumliche Effekte auf das Wahlverhalten. Die Hinzunahme dieser Parameter vermeidet die Missspezifikationen der empirischen Modelle; dennoch schreibe ich den bestimmten Konstanten λ_j keine umfassende inhaltliche Bedeutung zu und versteh sie nicht umstandslos als „Valenz“ oder als Indikator für die Qualität einer Partei oder ihrer Kandidaten. Bei den nicht-räumlichen Prädiktoren der Wahlentscheidung versammelt die Koeffizientenmatrix θ_{ij} Effektparameter für individualspezifische Variablen, die im Modell allein vor dem Hintergrund einer ausgewählten Referenzkategorie interpretierbar sind. Schließlich identifiziert ε_{ij} ein stochastisches Modell, und für die Verteilung der Fehlerterme wird eine Gumbel/Type-I-Extremwertverteilung unterstellt.

Bereits Grynaviski und Corrigan (2006, S. 394) haben auf „a dizzying combination of options“ bei der Modellierung von Wahlentscheidungen und der statistischen Abbildung der räumlichen Theorie des Wählens verwiesen. Dieser Beitrag konzentriert sich auf den Umgang der Wähler mit Risiko und Unsicherheit und kann zu einigen weiteren konzeptionellen Entscheidungen nur kurSORisch Auskunft geben:

Erstens kann an dieser Stelle kein ausführlicher Beitrag zur Diskussion um die angemessene Bestimmung von Parteipositionen geleistet werden. Parteialternativen werden durch die idiosynkratische Wahrnehmung jedes einzelnen Wählers und nicht durch ihre Mittelwerte oder irgendwelche übrigen „objektiven“ Positionsmaße platziert. Diese Strategie wird besonders durch das Argument gerechtfertigt, dass jeder einzelne Wähler seine ohnehin subjektive und nie vollständig „rationale“ Wahlentscheidung nur an denjenigen Informationen festmachen kann, über die er selbst verfügt.

Zweitens kann ich an dieser Stelle nicht zur Diskussion über die Vorzüge und/oder Nachteile des „Conditional Logit“-Modells gegenüber alternativen, komplexeren Techniken wie dem generalisierten Extremwertmodell oder dem alternativenspezifischen Probit-Modell beitragen (vgl. Alvarez und Nagler 1998). Als besonderer Vorzug der komplexeren Modelle wird häufig die Aufgabe der „independence of irrelevant alternatives“-Annahme genannt. Eine Reihe von einschlägigen Beiträgen zeigt jedoch, dass die IIA-Annahme selten empirisch relevant oder besonders restriktiv ist und die alternativen Modelle, insbesondere das multinomiale Probit-Modell, regelmäßig an Identifikationsproblemen leiden, die ihre praktische Verwendbarkeit sehr limitieren (vgl. Adams et al. 2005; Dow und Endersby 2004; Merrill und Adams 2002).

Drittens beschränke ich meine empirische Applikation auf eindimensionale politische Räume. Ohne Frage ist Politik in der modernen Massendemokratie stetig komplexer geworden, und der politische Wettbewerb ist häufig in zwei- oder mehrdimensionalen politischen Räumen organisiert. Dennoch hat eine Vielzahl neuerer Beiträge gezeigt, dass eine übergreifende ideologische Dimension, meist charakterisiert als Gegensatz von „links“ und „rechts“ oder „konservativ“ und „liberal“, weiterhin wesentliche Einzeldimensionen des politischen Wettbewerbs einschließt und, insbesondere im Kontext zunehmender politischer Polarisierung, die Struktur des politischen Wettbewerbs angemessen abbildet (Gabel und Huber 2000; Jesse 2012; Munger und Munger 2015). Die Konzentration auf einfache, eindimensionale Räume hat aber auch konzeptionelle und theoretische Gründe. Neben vielen anderen hat zum Beispiel Eguia (2013) eindringlich illustriert, dass die Charakterisierung politischer Präferenzen durch euklidische Verlustfunktionen in mehrdimensionalen Räumen oft problematisch und voraussetzungsvoll ist und insbesondere die unterschiedliche Salienz der räumlichen Dimensionen und die Frage nach ihrer Separierbarkeit eine Fülle weiterer Komplikationen in theoretische und statistische Modelle der Wahlentscheidung hineinträgt.

4.2 Empirische Befunde des reinen räumlichen Modells

Ich beginne die Vorstellung der empirischen Befunde mit einigen simplen Basismodellen. Die Daten sind dem kumulierten CSES-Datensatz entnommen, und ich schätze für jeden einzelnen der neunzig Wahlkontexte individuelle „Conditional Logit“-Modelle der Wahlentscheidung. Schlüsselvariable sind die räumlichen Distanzen von Wählern und Parteioptionen auf einer unidimensionalen Links-Rechts-Skala ($v_i; p_{i,j} \in [0, 10]$).

Das einfache Basismodell mit den Effektparametern α und β kann gleichermaßen durch konventionelle Maximum-Likelihood-Verfahren oder mit bayesianischer MCMC-Simulation ausgewertet werden, und für einfache Modelle erreichen beide Verfahren tendenziell identische Resultate. Bei der folgenden Verallgemeinerung hin zu komplexeren Modellen sind die Bayes-Verfahren jedoch deutlich flexibler und bieten zudem verlässlichere Angaben der statistischen Unsicherheit. Ich verwende in diesem Beitrag durchgehend nicht-informative, diffuse Priorverteilungen und simuliere die Posteriorverteilungen von α , β und den alternativenspezifischen Konstanten λ_j mit JAGS. Bei jedem der neunzig heterogenen Wahlkontakte benutze ich vier Markow-Ketten mit jeweils 100.000 Iterationen nach einem „burn in“ von 50.000 Iterationen. Eine Reihe von Spezifikationstests zeigt, dass jedes einzelne dieser parallel konstruierten „Conditional Logit“-Modelle das Datenmaterial angemessen abbildet. Konventionelle Geweke- und Heidelberger-Welch-Tests verweisen auf die Stationarität der Posteriorverteilungen. Ich speichere für die Beschreibung der Verteilungen nur jede zehnte Iteration („thinning“), um mit der vorhandenen Autokorrelation der Markow-Ketten angemessen umzugehen. Schließlich simulierte ich die Modelle effizienter, indem ich Kernparameter wie α und β blocke, also aus einer multivariaten Normalverteilung ziehe (theoretische und praktische Aspekte der MCMC-Simulation werden genauer beschrieben bei Gelman et al. 2013; Jackman 2009).

Abb. 2 zeigt Zusammenfassungen der Posteriorverteilungen des Salienzparameters (α ; x-Achse) und des Formparameters (β ; y-Achse). Beide Teilbilder bieten im wesentlichen identische Informationen: Das obere summiert bayesianische Punktschätzer und Konfidenzintervalle; das untere ersetzt diese Informationen durch die konkrete Benennung einzelner Umfragesegmente, sodass die meisten Wahlkontakte identifizierbar werden.

Die horizontale Dimension von Abb. 2 bestätigt eindeutig die Aussagekraft der räumlichen Theorie der Wahlentscheidung über eine Vielzahl heterogener Wahlen und Kontexte des kumulierten CSES-Datensatzes hinweg. Bei jeder einzelnen Wahl ist die Nähe oder Distanz von Wählern und Parteien im politischen

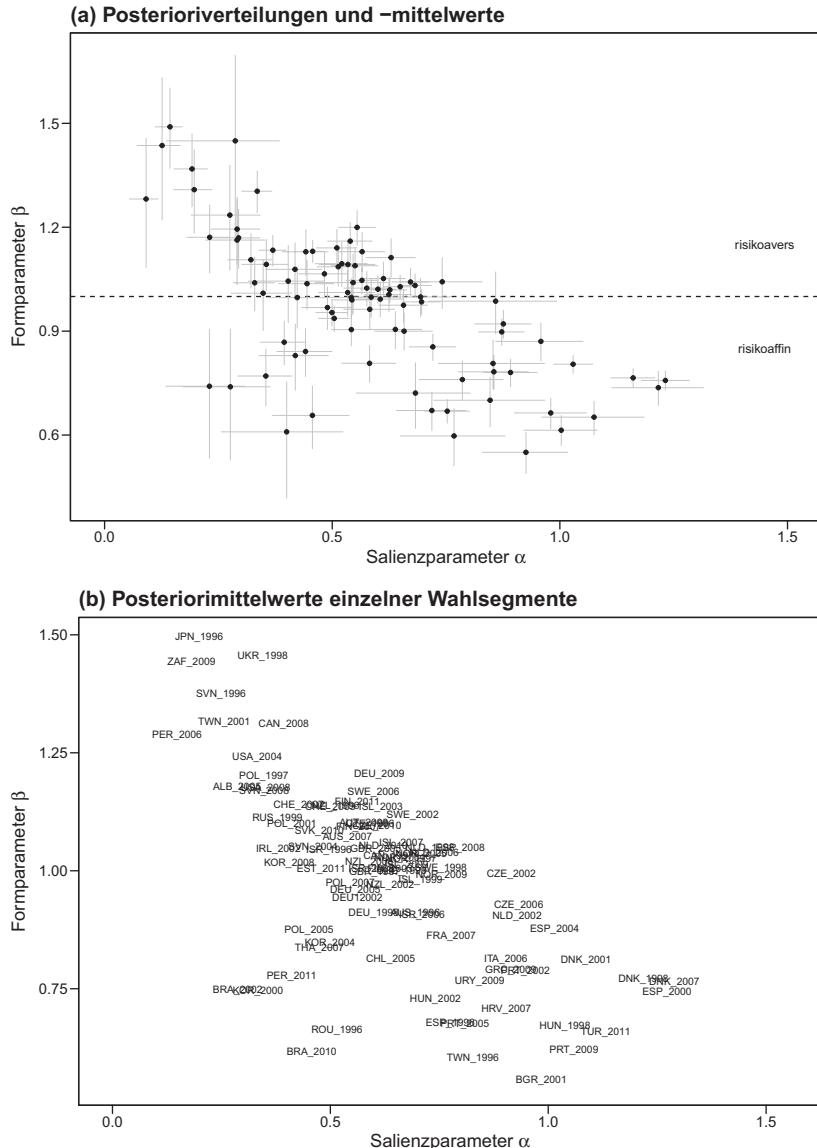


Abb. 2 Schätzer für Salienz- und Formparameter ($\hat{\alpha}$ und $\hat{\beta}$)

Raum eng und substanzial bedeutsam mit der Wahlentscheidung verbunden ($\hat{\alpha} > 1$). Die simulierten Modellparameter zeigen eindeutig, dass die räumliche Theorie des Wählens über ganz unterschiedliche soziale, politische und institutio-nelle Kontexte hinweg wesentliche Einsichten in die Bewertung politischer Parteien und in die Wahlentscheidung eröffnet. Nicht nur in abstrakten theoretischen Modellen, sondern auch in konkreten empirischen und statistischen Analysen bevorzugen Wähler diejenigen Parteien, die ihren eigenen Idealpunkten möglichst genau entsprechen gegenüber denjenigen Plattformen, die sie weiter entfernt im politischen Spektrum verorten.

Während die räumliche Nähe auf der Links-Rechts-Dimension über alle Kon-texte hinweg einen Effekt auf die Wahlentscheidung hat, unterscheiden sich die Stärke dieses Effekts und die Salienz des ideologischen Konflikts von links und rechts erheblich. Der Einfluss der räumlichen Nähe von Wählern und Parteialter-nativen nimmt ab, wenn politische Parteien und Parteiensysteme sich im stetigen Umbruch befinden (zum Beispiel in Albanien, Brasilien oder Peru). Die Wir-kungen ideologischer Nähe auf der Links-Rechts-Dimension sind dagegen sehr viel ausgeprägter wenn der politische Wettbewerb hinreichend institutionalisiert, strukturiert oder auch polarisiert ist (etwa in Dänemark, Spanien oder Ungarn.) In diesem kurzen Beitrag kann ich mögliche Ursachen nur andeuten: Auf der einen Seite erzeugen die Stabilisierung und Institutionalisierung notwendige Funktions-bedingungen für die Entstehung programmatisch strukturierter Verbindungen zwischen Wählern und politischen Parteien. Zudem charakterisiert der Gegensatz von Links und Rechts den politischen Wettbewerb in verschiedenen Vergleichs-staaten unterschiedlich gut. Besonders im Kontext schroffer Polarisierung wird häufig die Dimensionalität des politischen Raums reduziert und die Struktur des politischen Wettbewerbs korrespondiert dann regelmäßig besser mit einem über-greifenden, eindimensionalen Links-Rechts-Schema.

Eine systematische Metaanalyse unterstreicht die Stabilität der empirisch bestimmen Parameter über wiederholte Wahlen innerhalb desselben Ver-gleichsstaats hinweg: Die Inspektion der Salienzparameter $\hat{\alpha}$ verweist auf sehr hohe Autokorrelation unter Beobachtungen aus denselben Vergleichsstaaten ($p[\hat{\alpha}] = 0.72$). Deshalb liegt nahe, dass die Bedeutung ideologischer Nähe oder Distanz für die Wahlentscheidung auch eine systemische Variable ist, die von politischen Institutionen wie Parteiensystemen, Konfliktstrukturen oder institutio-nellen Regeln mindestens mitbestimmt wird.

Die generelle Aussagekraft des räumlichen Modells ist eine notwendige Bedingung für das Kernargument dieses Beitrags. Das grundlegende Interesse zielt ja darauf, die Form der Nutzenfunktionen innerhalb verschiedener Elek-torate empirisch zu ermitteln und diese Befunde systematisch zu vergleichen,

um so Aufschlüsse über die (durchschnittliche) Risikoorientierung der Wähler zu gewinnen. Schätzer für den Formparameter $\hat{\beta}$ sind in der vertikalen Dimension von Abb. 2 abgetragen. Diese Befunde, jeweils über neunzig einzelne, unabhängig voneinander geschätzte „Conditional Logit“-Modelle hinweg, illustrieren eindringlich, dass in der überwiegenden Mehrzahl der untersuchten Wahlen die Wähler eine risikoneutrale Position einnehmen ($\hat{\beta} \sim 1$). Nicht in einem einzigen Kontext dagegen entsprach der empirisch bestimmte Formparameter der strikten Annahme risikoaverser Wähler im „Neo-Downsianischen“ Standardmodell, die bei der Verwendung quadratischer Verlustfunktionen unterstellt wird ($\hat{\beta} = 2$). Dennoch sind bei einer Reihe von Parlamentswahlen, zum Beispiel in Japan, Südafrika oder in der Ukraine, die Elektorate tendenziell risikoavers. In anderen Kontexten, zum Beispiel in Bulgarien, Rumänien, oder Ungarn, verläuft die Nutzenfunktion des Elektorats dagegen mindestens teilweise konvex und die Elektorate sind somit tendenziell risikoaffin.

Wie zuvor verweist eine Metaanalyse auf substanzelle Autokorrelation ($p[\hat{\beta}] = 0.48$) und unterstreicht deshalb auf die Stabilität dieser Befunde bei wiederholten Wahlen innerhalb derselben Vergleichsstaaten. Damit liegt nahe, dass auch die Risikoorientierung der Elektorate durch langfristige, stabile institutionelle Rahmenbedingungen mindestens mitgeprägt wird.

Diese Befunde hinterfragen einige wesentliche Prämissen, die der Spezifikation der räumlichen Theorie des Wählens regelmäßig zugrunde liegen. Über ein weites Spektrum unterschiedlicher Vergleichsstaaten und heterogener Kontexte hinweg sind Elektorate und Wähler nicht generell risikoavers, sondern die Mehrheit der Wähler verhält sich risikoneutral. Die Annahme konsistent risikoaverser Wähler, die direkt durch die Benutzung quadratischer Verlustfunktionen abgebildet wird, ist deshalb viel zu restriktiv, und sie wird durch die umfangreichen Daten, auf denen diese Analysen basieren, nirgendwo gedeckt.

Am Schluss dieses Abschnitts möchte ich auf einige potenzielle Probleme dieses Modells verweisen. Diese Bedenken betreffen besonders die mögliche Modellabhängigkeit der Resultate. Ganz unabhängig davon, ob bestimmte Formen der Nutzenfunktionen *a priori* definiert (wie bei Gynaviski and Corrigan, 2006 oder Singh 2014) oder empirisch geschätzt werden (wie in diesem Beitrag) beeinflusst die definierte oder geschätzte numerische Größe des Formparameters das Niveau des Salienzparameters. Die theoretische Spezifizierung der Nutzenfunktionen und die einfache Inspektion der empirischen Modellparameter in Abb. 2 verweisen nachdrücklich darauf, dass $\hat{\alpha}$ und $\hat{\beta}$ invers miteinander verbunden sind. Hohe Schätzer für die Salienz der ideologischen Dimension $\hat{\alpha}$ korrelieren recht deutlich mit geringeren Werten für den Formparameter $\hat{\beta}$.

Freilich sollte auf der Grundlage der hier benutzen Daten und mit den Bedenken über die mögliche Modellabhängigkeit dieser Befunde daraus nicht gefolgert werden, dass hohe Salienz und risikoaffine Wählerdispositionen einander kausal und theoretisch bedingen.

4.3 Berücksichtigung und Modellierung nicht-räumlicher Kontrollvariablen

Nur wenige empirische Analysen begründen die Wahlentscheidung *allein* mit der Bewertung ideologischer oder programmatischer Passfähigkeit eines Wählers mit den verschiedenen Parteialternativen. Stattdessen schlagen die meisten neueren Beiträge „vereinigte“ Modelle vor, die simultan räumliche, nicht-räumliche und valenzbasierte Motive der Wahlentscheidung aufnehmen (Adams et al. 2005; Jessee 2012; Merrill und Grofman 1999; Schofield und Sened 2006). Das umfassende Datenmaterial des CSES-Projekts enthält eine Reihe von nicht-räumlichen Variablen, die die Wahlentscheidung mitbeeinflussen *und* die gleichermaßen mit der Nähe oder Distanz im räumlichen Modell verbunden sind, zum Beispiel die Art und Stärke der Parteidentifikation, die Zufriedenheit mit dem demokratischen Prozess, den Grad an politische Information eines Befragten und grundlegende demografische Kategorien wie Alter, Geschlecht und Bildungsniveau.

Dieser Beitrag interessiert sich weniger für die direkten Wirkungen dieser Variablen auf das Wahlverhalten, sondern bestimmt, ob die Aufnahme dieser Kontrollvariablen die Befunde einfacher räumlicher Modelle erhält oder modifiziert. Die Prädiktormatrix D_i bildet die Effekte einer Reihe von nicht-räumlichen Einflüssen ab:

1. Die Parteidentifikation, bestimmt als die Angabe eines langfristigen Näheverhältnisses zu einer konkreten politischen Partei, ist per se ein individualspezifisches Charakteristikum. Für die Aufnahme in das statistische Modell wird diese Variable in eine Reihe binärer Dummy-Variablen heruntergebrochen, die anzeigen, ob sich ein Wähler i mit einer bestimmten Parteialternative j identifiziert oder nicht.
2. Als eine weitere nicht-räumliche Kontrollvariable geht die Zufriedenheit mit dem politischen Prozess in die Modelle ein. Dieser Indikator wird auf einer einfachen Vierpunktskala erhoben und als ein Proxy für die Zufriedenheit mit der amtierenden Regierung verwendet. Konkrete Bewertungen der

Regierungsarbeit werden leider nur in der zweiten und dritten Welle der CSES-Umfragen erhoben.

3. Jede Analyse von Wahlentscheidung, die Risiko und Unsicherheit angemessen berücksichtigen möchte, benötigt Daten zur politischen Informiertheit einzelner Wähler. Die CSES-Questionnaires enthalten drei binäre Wissensfragen, die zu einer Skala politischer Informiertheit aufsummiert werden.
4. Schließlich gehen als demografische Kontrollvariablen das Alter, das Geschlecht und das Bildungsniveau der Befragten in die jeweiligen Modelle ein.

Abb. 3 vergleicht die Posterioriverteilungen für den Salienzparameter ($\hat{\alpha}$; linke Bildseite) und den Formparameter ($\hat{\beta}$; rechte Bildseite). Auf der x -Achse sind jeweils die Schätzer für einfache räumliche Modelle abgetragen, und die y -Achse kontrastiert die Parameter für reichhaltigere, vereinigte Modelle mit den spezifizierten Kontrollvariablen. Bereits eine oberflächliche Inspektion beider Modelle zeigt deutlich, dass die Hinzunahme nicht-räumlicher Kontrollvariablen die Parameter des Modells kaum modifiziert. Die zusätzliche Berücksichtigung signifikanter Einflüsse wie Parteidentifikation, Demokratiezufriedenheit, politische Information und demografischen Basisvariablen verringert oder erhöht nicht die generelle Aussagekraft der räumlichen Modellkomponente ($\hat{\alpha}$), sie hat keinen Einfluss auf die Entstehung und Veränderung konkaver oder konvexer Nutzenfunktionen und sie betrifft nicht das hiervon abgeleitete Risikoverhalten der Wähler ($\hat{\beta}$).

4.4 Berücksichtigung und Modellierung unsicherer Parteipositionen

Ich komplettiere nun das statistische Modell und füge die Unsicherheit von Wählern über die politischen Standorte der politischen Parteien als messbare und gemessene Variable hinzu. Dabei sind valide Bestimmungen individueller Unsicherheit mit den durch weitgehend standardisierte Wahlforschungsprojekte wie CSES bereitgestellten Daten oft konzeptionell schwierig und innerhalb der Wahlforschung umstritten (vgl. die detaillierte Diskussion und Bewertung alternativer Unsicherheitsmaße bei Alvarez 1998).

Bei der Operationalisierung von individueller Unsicherheit über die Positionen einzelner Parteien dominieren im Wesentlichen zwei, teils verschiedene, teils komplementäre Perspektiven: Bartels (1986) argumentiert, dass im Grunde alle Wähler unsicher seien über die Positionen aller Parteien, dass sie sich jedoch teils

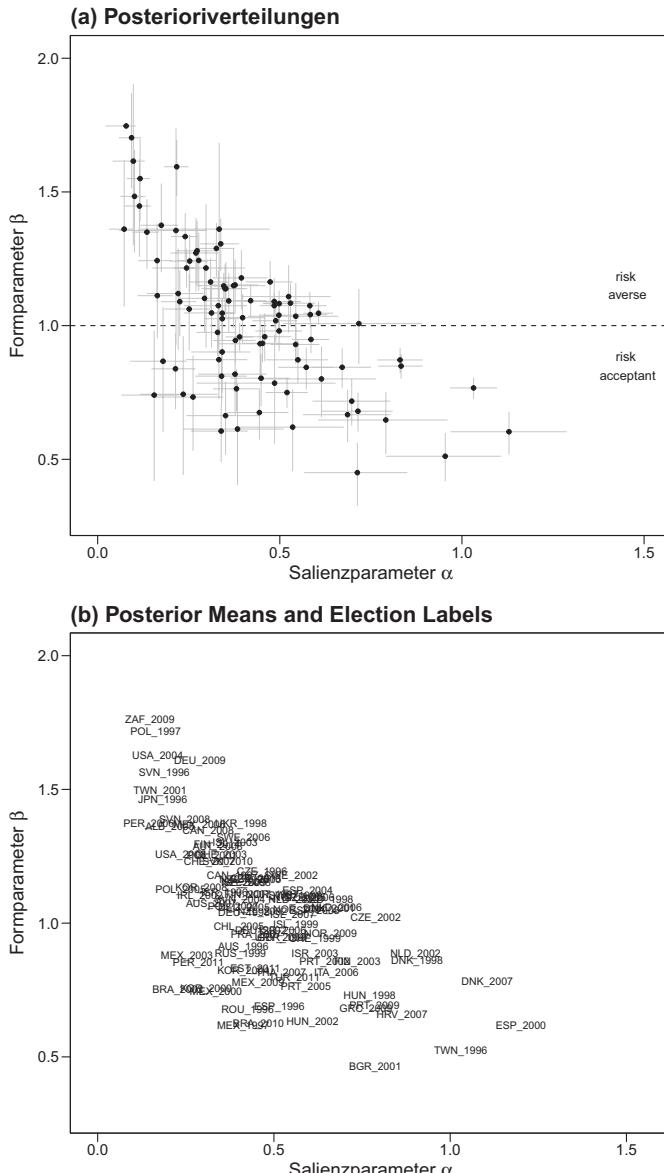


Abb. 3 Vergleich von einfachen und gesättigten Modellen ($\hat{\alpha}$ and $\hat{\beta}$)

erheblich im Grad dieser Unsicherheit unterscheiden. Wenn jedoch die Unsicherheit eines Wählers i über die Platzierung einer Partei j ein bestimmtes Niveau überschreite, werde dieser Befragte eine Verortung dieser Partei auf einer ideo-logischen oder politischen Skala ablehnen. Unsicherheit kann in diesem Sinne nicht direkt beobachtet werden, sondern Muster an Antwortverweigerungen werden als Maßstab für die Unsicherheit der Wähler benutzt. Im zweiten Schritt formuliert Bartels (1986) ein statistisches Modell, das fehlende Antworten durch eine Reihe von Erklärungsfaktoren abbildet und zum Beispiel Bildung, politische Informiertheit, politisches Interesse und die Verarbeitung von Kampagnen-informationen und Parteiwerbung berücksichtigt. Die mit den empirischen Modellparametern vorhergesagte Wahrscheinlichkeit einer Antwortverweigerung wird dann als Maßstab und Instrument von Sicherheit oder Unsicherheit über die jeweilige Parteiposition bewertet.

Anstatt dieser inferenzbasierten Modellierungsstrategien unterstreicht Alvarez (1998) die Vorzüge direkt beobachteter Indikatoren. Er unterstellt, dass jede Platzierung einer Partei durch einen Wähler $p_{i,j}$ aus ihrem realen, „wahren“ Standort p_j und aus einem wählerv- und parteispezifischen Fehlerterm $\delta_{i,j}$ zusammengesetzt sei: $p_{i,j} = p_j + \delta_{i,j}$. Die Unsicherheit eines Wählers i bei der Verortung von Partei j wird deshalb durch die Differenz individualspezifischer und „wahrer“ oder durchschnittlicher Parteipositionen bestimmt: $\delta_{i,j} = |p_{i,j} - p_j|$. Falls jedoch ein Befragter für eine bestimmte Partei keine Position angibt, definiere ich für den Indikator die maximal mögliche Distanz, also bei einer Elfpunkteskala gilt $\delta_{i,j} = 11$. Bei der Messung politischer Unsicherheit folge ich diesem Vorschlag von Alvarez (1998), weil die Qualität des von Bartels (1986) vorgeschlagenen Instruments in empirischen Applikationen oft nicht sehr hoch ist, sodass ich direkte gegenüber inferentiellen Indikatoren politischer Unsicherheit vorziehe.

Der Unsicherheitsindikator $\delta_{i,j}$ nimmt verschiedene Werte ein, die für jeden Wähler i und jede Partei j separat bestimmt werden. Konkrete Werte von $\delta_{i,j}$ reichen von 0, bei $p_{i,j} = \bar{p}_j$, bis hin zu 11, wenn $p_{i,j}$ und \bar{p}_j an den gegenüber liegenden Extrempositionen der Skala positioniert sind oder wenn der Wähler i keine Platzierung von Partei j vornehmen kann oder möchte. Durch diese Regeln zum Umgang mit fehlenden Platzierungen sind die Werte von $\delta_{i,j}$ stets bimodal ver-teilt. Spezifische Werte des Indikators $\delta_{i,j}$ schwanken recht deutlich, und zwar sowohl innerhalb einzelner Segmente des CSES-Datensatzes und über diese Segmente hinweg. Unter den neunzig Umfragesegmenten in der Analyse waren sich zum Beispiel die Befragten in Tschechien 2002 am sichersten über die Standorte der Parteien ($\bar{\delta}_{i,j} = 1,26$), in Italien 2006 waren sie dagegen besonders unsicher ($\bar{\delta}_{i,j} = 4,97$).

Das statistische Modell korrespondiert genau mit dem theoretischen Konzept von Unsicherheit, das eingangs dargestellt wurde. Es gründet im Wesentlichen auf einer Modellbildungsstrategie, die Bartels (1986) entworfen, Berinsky and Lewis (2007) durch die explizite Modellierung verschiedener Risikoorientierungen weiterentwickelt und shikano and Behnke (2009) repliziert und für Mehrparteiensysteme verallgemeinert haben. Wie zuvor wird die Auswahl unter mehr als zwei Parteialternativen mit einem bayesianischen „Conditional Logit“-Modell abgebildet. Anstatt jedoch jede einzelne Partei durch einfache Punktschätzer im politischen Raum zu verorten, charakterisiere ich die Wahrnehmung unsicherer Parteien durch eine Wahrscheinlichkeitsverteilung über ihrer jeweiligen Position. Dabei wird bei jeder Iteration des MCMC-Prozesses ein neuer Wert für die Parteiposition aus dieser Verteilung gezogen und für die Schätzung der Modellparameter verwendet.

Im nächsten Schritt verbinde ich die beiden eben vorgestellten Ideen, die direkte Messung der Wählerunsicherheit und die Charakterisierung als unsicher wahrgenommener Parteipositionen durch Wahrscheinlichkeitsverteilungen. Die Grundannahme lautet, dass beide Größen proportional zueinander sind: Je höher die Unsicherheit eines Wählers i über die Position von Partei j , desto größer ist die Standardabweichung einer Normalverteilung über dieser Position: $\sigma(p_{i,j}) \propto \delta_{i,j}$. Bartels (1986, 717) unterstellt dabei, der Unsicherheitsindikator sei entworfen „to reflect variances of candidate perceptions up to an (unknown) positive scale factor“. Genau deshalb müssen die empirisch bestimmten Werte von $\delta_{i,j}$ durch einen weiteren Modellparameter γ reskaliert werden, und das Produkt $\gamma\sigma[p_{i,j}]$ wird so als die tatsächliche empirische Variation der Parteipositionen interpretierbar:

$$\Theta_{i,j} = N(p_{i,j}, \sigma[p_{i,j}]) = N(p_{i,j}, \gamma\delta_{i,j}) \quad (5)$$

Schließlich ersetze ich die fixierten Parteipositionen $p_{i,j}$ durch die Wahrscheinlichkeitsverteilungen $\Theta_{i,j}$:

$$u(v, \Theta_{i,j}) = -\alpha|v_i - \Theta_{i,j}|^\beta + \lambda_j c_j + \varepsilon_{i,j}; Pr(v_i = j) = \frac{\exp[u(v, \Theta_{i,j})]}{\sum_{j=1}^J \exp[u(v, \Theta_{i,j})]} \quad (6)$$

Bei diesem komplexeren Modell ist die technische Bestimmung der Modellparameter durch MCMC-Simulationen nicht einfach, weil die einzelnen Markow-Ketten häufig eine hohe Autokorrelation aufweisen und bei der Parameterschätzung nur langsam mischen. Der MCMC-Prozess simuliert nicht nur die einzelnen Modellparameter, sondern auch die Parteipositionen müssen jeweils aus Unsicherheitsverteilung $\Theta_{i,j} = N(p_{i,j}, \gamma\delta_{i,j})$ gezogen werden. Deshalb habe ich für die

Simulation dieser komplexeren Modelle in jedem der neunzig Wahlsegmente vier Ketten und ein „burn in“ von 100.000 Iterationen vorgesehen und weitere 200.000 Iterationen für die Charakterisierung der Posterioriverteilungen gespeichert. (Ein Ausschnitt des für die MCMC-Simulationen benutzten JAGS-Codes ist in Anhang B dokumentiert.).

Abb. 4 fasst die wesentlichen Modellbefunde zusammen und charakterisiert die Posterioriverteilungen für die Salienz- ($\hat{\alpha}$), Form- ($\hat{\beta}$) und Unsicherheitsparameter ($\hat{\gamma}$). Die linke Spalte (a) charakterisiert zunächst den Salienzparameter $\hat{\alpha}$. Die geschätzten Parameter sind über neunzig heterogene Wahlumgebungen hinweg stets positiv, und diese Befunde belegen, dass, auch wenn die Unsicherheit von Parteiplatzierungen explizit modelliert wird, die räumliche Nähe auf der ideologischen Links-Rechts-Dimension weiterhin einen erheblichen Einfluss auf die Bewertung politischer Parteien ausübt und die Wahlentscheidung mitbestimmt. Über die einzelnen Kontexte hinweg variiert dieser Einfluss jedoch oft erheblich.

Diese Resultate bestärken einerseits die bereits vorgestellten und diskutierten Befunde einfacher räumlicher Modelle. Während die Verteilungen der Parameter $\hat{\alpha}$ und $\hat{\beta}$ über die einzelnen Wahlsegmente hinweg beinahe unverändert bleiben, modifiziert die systematische Berücksichtigung unsicherer Parteipositionen die Befunde für einzelne Umfragesegmente des CSES-Datensatzes oft deutlich. Eine Konsequenz ist, dass die Modellabhängigkeit der Parameter im einfachen räumlichen Modell ($-\alpha \sim \beta$) in den komplexeren Modellen nicht reproduziert wird. Benutzt man die Wählerunsicherheit δ_{ij} , um die Verteilungen unsicher wahrgenommener Parteipositionen θ_{ij} zu charakterisieren, sind die Salienz- und Formparameter empirisch unabhängig voneinander.

Die mittlere Spalte (b) in Abb. 4 fasst die Posterioriverteilungen des Form- oder Risikoparameters $\hat{\beta}$ zusammen. Die systematische Berücksichtigung der Wählerunsicherheit ändert zunächst nur wenig an seiner empirische bestimmten Bandbreite über die heterogenen Wahlkontakte hinweg. Die Mittelwerte der Posterioriverteilungen liegen meist recht eng bei $\hat{\beta} = 1$, unterstreichen die empirische Angemessenheit linearer Distanzen und zeigen eine klare Tendenz hin zu risiko-neutralen Haltungen der meisten Elektorate an. Dagegen werden einige Standardannahmen des „Neo-Downsianischen“ Modells, die unbedingte Bevorzugung quadratischer, konkaver Nutzenfunktionen und die Unterstellung einer strikt risikoaversen Perspektive der Wähler und Elektorate, durch empirische Modellbefunde weiterhin beinahe nirgendwo gestützt.

Schließlich stellt die rechte Spalte (c) in Abb. 4 die Posterioriverteilungen des Unsicherheitsparameters γ dar. Stets positive Schätzer und beinahe durchgehend

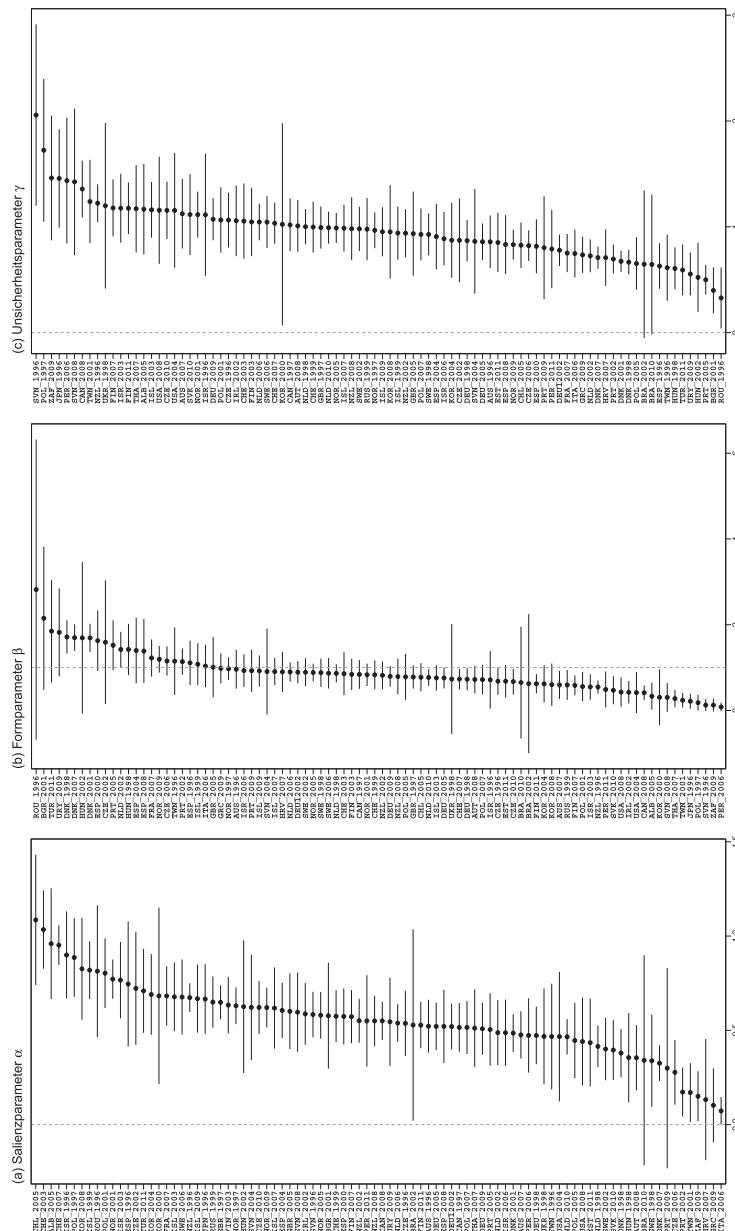


Abb. 4 Posterioriverteilungen der Salienz-, Form- und Unsicherheitsparameter ($\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$)

positive Konfidenzintervalle des Unsicherheitsparameters verweisen auf den Einfluss unsicher wahrgenommener Parteipositionen auf die Wahlentscheidung. Höhere Parameterschätzer für $\hat{\gamma}$ korrespondieren mit weiteren Unsicherheitsverteilungen über der jeweiligen Parteiposition $p_{i,j}$ und verweisen auf die gesteigerte Bedeutung von Unsicherheit für die Bewertung von Parteialternativen und die Wahlentscheidung.

5 Fazit und Ausblick

Theoretische und empirische Beiträge zur räumlichen Theorie des Wählers haben oft einfach unterstellt, wie Wähler mit Risiko und Unsicherheit umgehen. Dieser Beitrag geht dagegen einen anderen Weg und bestimmt die Salienz räumlicher Nutzenterme und die Logik, mit der Wähler räumliche Distanzen in Nutzenterme verwandeln, in einer empirisch-vergleichenden Analyse über eine Fülle heterogener Kontexte hinweg. Diese Vogelperspektive auf die Risikoorientierung gesamter Elektorate erschließt mindestens zwei wichtige Einsichten:

1. Die standardisierte Annahme konkaver, quadratischer Nutzenfunktionen ist für die valide Modellspezifikation zu restriktiv. Die empirischen Resultate für die meisten der heterogenen Kontexte des CSES-Trendfiles widersprechen dieser Annahme ausdrücklich und verweisen stattdessen tendenziell auf die Angemessenheit und Nützlichkeit linearer Verlustfunktionen für die meisten Entscheidungssituationen.
2. Weil die Form der Verlustfunktionen bestimmt, wie Wähler geometrische Distanzen in die Nutzenterme des räumlichen Modells verwandeln, betrifft sie direkt die Risikoorientierung der Wähler. Die empirischen Resultate widersprechen der klassischen Annahme strikt konkaver und risikoaverser Präferenzen. Über heterogene Kontexte hinweg verfolgen Wähler stattdessen tendenziell risikoneutrale Haltungen, und in einer Minderheit der Wahlsegmente handeln Wähler sogar risikoaffin.

Die inhaltliche Bewertung dieser Befunde sollte stets mit großer Vorsicht erfolgen, und sie wird deshalb mit einigen Qualifikationen vorgenommen. Zunächst, das wurde bereits diskutiert, erschließen die hier vorgestellten Modelle die Risikoorientierung ganzer *Elektorate*, nicht einzelner Wähler. Nun kann aber nicht ohne Weiteres unterstellt werden, dass alle Befragten oder gar alle Wähler bei allen Parteien, die in einem Land zur Wahl stehen, dieselben Risikobewertungen anlegen. Diese Verlustfunktionen spezifisch für Gruppen von Wählern und für einzelne Parteien zu bestimmen, bleibt deshalb ein wichtiges Desiderat der Forschung.

Nicht nur die fehlende Möglichkeit, konkrete Risikoorientierungen einzelner Wähler zu bestimmen, zu modellieren und für empirische Analysen aufzugreifen, spricht für die mindestens komplementäre Hinzunahme experimenteller Studien. Die Verwendung einfacher Umfragedaten, auch wenn das hier aus einer empirisch gesättigten und strikt vergleichenden Perspektive geschieht, kann nur Aufschlüsse über empirische Assoziationen von Unsicherheit und Risikoorientierung bieten. In welche Richtung die Effekte genau weisen, was Eplanans und was Explanandum ist, bleibt jedoch nur teilweise bestimmt. Formale Modelle definieren häufig die Risikoorientierung der Wähler als eine strikt exogene Größe und postulieren, dass Wählerverhalten und Parteistrategie einfach darauf reagieren. Vice versa kann jedoch auch argumentiert werden, dass politische Unsicherheit, Informationsdefizite, widersprüchliche Informationen am Ausgangspunkt der Kausalketten stehen und ihrerseits Risikokalkulationen rationaler und nicht-rationaler Wähler erst begründen.

Zudem ist nicht nur in der formalen Literatur weiterhin umstritten, ob die hier vorgestellten Befunde umstandslos als stringente Belege für risikoneutrales oder sogar risikoaffines Verhalten bewertet werden können. Bei der formalen Modellspezifikation verweist zum Beispiel Eguia (2013) eindringlich auf Probleme euklidischer Nutzenfunktionen in mehrdimensionalen politischen Räumen, die Gewichtung unterschiedlicher Wettbewerbsdimensionen und die analytische Separierbarkeit dieser Nutzenkomponenten. Schließlich wenden Shikano und Behnke (2009) ein, dass selbst belegbar konvexe Verlustfunktionen nicht unbedingt und ausschließlich auf risikoaffines Verhalten verweisen. Stattdessen könnten konvexe Verlustfunktionen auch für die mangelnde Bereitschaft sprechen, politische Positionen angemessen zu differenzieren und Güterabwägungen vorzunehmen. Sie wären damit eher ein empirischer Hinweis auf die Rigidität der Wähler als auf ihre Risikobereitschaft.

Schließlich kann der Effekt von Unsicherheit und Risikoorientierung in räumlichen Modellen nicht mit der exklusiven Perspektive auf einzelne Wähler und gesamte Elektorate untersucht werden. Eine wirklich komplette Analyse müsste auch die „Angebotsseite“ politischer Parteien angemessen berücksichtigen und bestimmen, wann und warum sie im politischen Wettbewerb unklare ideologische und/oder programmatische Positionen anbieten.

Danksagung Ich danke Susumu Shikano für Rat und Hilfe bei der Spezifikation der statistischen Modelle, und ich danke den beiden anonymen Gutachtern für die ihre sehr wertvollen Kommentare und Vorschläge. Der Beitrag entstand im Rahmen des von der Fritz-Thyssen-Stiftung geförderten Projekts „Lost in Space? The Emptiness of the Center and Centrifugal Determinants of Vote Choice and Party Competition in EP Elections“ (Az. 10.17.1.039PO).

Appendix

Das Datenmaterial des CSES-Projekts

Umfragesegmente des kumulierten „CSES-Trendfile“

Die empirischen Analysen in diesem Beitrag stützen sich ausschließlich auf das reichhaltige Datenmaterial des CSES-Projekts. Ich verwende den „CSES Harmonized Trend File“, der länder- und wahlspezifischen Umfragesegmente von 1996 bis 2011 standardisiert und für die empirische Datenanalyse aufbereiten. Dieser integrierte Datensatz wurde am Wissenschaftszentrum Berlin von Heiko Giebler, Josephine Lichteblau, Antonia May, Reinhold Melcher, Aiko Wagner und Bernhard Weßels zusammengestellt. Daten und Dokumentation sind auf den Webseiten des Projekts verfügbar (<http://cses.org/datacenter/trendfile/trendfile.htm>; Stand: 15. Dezember 2015).

Diese Aufbereitung ersten drei Wellen des CSES-Projekts integriert und harmonisiert diejenigen Variablen, die wiederholt, also in mindestens zwei Wellen, abgefragt wurden. Insgesamt vereinigt der CSES-Trendfile Datenmaterial aus 129 einzelnen Umfragesegmenten, die im Kontext von Wahlen zu nationalen Parlamenten und zur nationalen Präsidentschaft abgehalten wurden. Um die Vergleichbarkeit der untersuchten Wahlbefragungen zu gewährleisten, habe ich allein Wahlen zu nationalen Parlamenten (bei bikameralen Systemen zu den jeweiligen Unterhäusern) ausgewählt, die stärker personalisierten Präsidentschaftswahlen jedoch ausgeschlossen. Auch einige weitere Segmente konnten ich nicht berücksichtigen: Teils fehlten einige Schlüsselvariablen der Modelle, teils waren die Daten fehlerhaft oder inkonsistent, und einige Umfragesegmente waren so klein, dass keine sinnvollen Inferenzen möglich waren. Mit diesen Kriterien kann die Analyse nur neunzig der ursprünglich 128 Wahlsegmente aus insgesamt 44 verschiedenen Staaten aufnehmen:

Albanien (2005), Australien (1996, 2007), Bulgarien (2001), Brasilien (2002, 2010), Chile (2005), Deutschland (1998, 2002, 2005, 2009), Dänemark (1998, 2001, 2008), Estland (2011), Finnland (2003, 2007, 2011), Frankreich (2007), Großbritannien (1997, 2005), Griechenland (2009), Ungarn (1998, 2002), Irland (2002), Island (1999, 2003, 2007, 2009), Israel (2003, 2006), Italien (2006), Japan (1996), Kanada (1997, 2007), Korea (2000, 2004, 2008), die Niederlande (1998, 2002, 2006, 2010), Norwegen (1997, 2001, 2005, 2009), Kroatien (2007), Neuseeland (1996, 2002, 2008), Österreich (2008), Peru (2006, 2011), Polen (1997, 2001, 2005, 2007), Portugal (2002, 2005), Rumänien (1996), Russland (1999), die Schweiz (1999, 2003, 2007), die Slowakei (2010), Slowenien (1996,

2004, 2008), Schweden (1998, 2002), Spanien (1996, 2000, 2004, 2008), Südafrika (2009), Thailand (2007), die Türkei (2011), Taiwan (1996, 2001), die Tschechische Republik (1996, 2002, 2006, 2010), die Ukraine (1998), Uruguay (2009) und die USA (2004, 2008).

Auswahl und Operationalisierung von Schlüssel- und Kontrollvariablen werden im Text knapp vorgestellt und begründet. Genaue Information zur Formulierung, Erhebung und Aufbereitung der einzelnen Indikatoren stehen kumuliert auf der Seite des CSES-Trendfiles und individuell auf den jeweiligen Seiten der drei benutzten CSES-Wellen bereit (<http://cses.org/>). Komplette Replikationsarchive stelle ich auf Anfrage gern zur Verfügung.

JAGS Code für die Schätzung des Unsicherheitsmodells

```
model{
for(i in 1:N_V){
  for(j in 1:N_P){
    mu[i,j] <- beta[1,j] - alpha[1] * (abs(lr_i[i]-lr_ij.dist[i,j]))^alpha[2]
    emu[i,j] <- exp(mu[i,j])
    p[i,j] <- emu[i,j]/sum(emu[i,1:N_P])

    prec.lr[i,j] <- pow(unc_ij[i,j] * gamma, -2)
    lr_ij.dist[i,j] ~ dnorm(lr_ij[i,j], prec.lr[i,j])
  }
  vote[i] ~ dcat(p[i,1:N_P])
}
# Priors;
# ALPHA;
alpha[1:N_ALPHA] ~ dmnorm(a0, A0)
# BETA;
# identifying restriction;
for(i in 1:N_BETA){
  beta[i,1] <- 0
}
for(i in 2:N_P){
  beta[1:N_BETA,i] ~ dmnorm(b0,B0)
}
# GAMMA;
gamma ~ dunif(0,3)
}
```

Literatur

- Adams, J., Merrill, S., and Grofman, B. (2005). *A Unified Theory of Party Competition. A Cross-National Analysis Integrating Spatial and Behavioral Factors*. Cambridge University Press, Cambridge.
- Aldrich, J. H. and McKelvey, R. D. (1977). A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections. *The American Political Science Review*, 71(1):111–130.
- Alvarez, R. M. (1998). *Information and Elections*. University of Michigan Press, Ann Arbor.
- Alvarez, R. M. and Nagler, J. (1998). When Politics and Models Collide: Estimating Models of Multiparty Elections. *American Journal of Political Science*, 42(1):55–96.
- Bartels, L. M. (1986). Issue Voting Under Uncertainty: An Empirical Test. *American Journal of Political Science*, 30(4):709–728.
- Berinsky, A. J. and Lewis, J. B. (2007). An Estimate of Risk Aversion in the U.S. Electorate. *Quarterly Journal of Political Science*, 2(2):139–154.
- Brady, H. E. and Ansolabehere, S. (1989). The Nature of Utility Functions in Mass Publics. *The American Political Science Review*, 83(1):143–163.
- Carroll, R., Lewis, J. B., Lo, J., Poole, K. T., and Rosenthal, H. (2013). The Structure of Utility in Spatial Models of Voting. *American Journal of Political Science*, 57(4):1008–1028.
- Davis, O. A., Hinich, M. J., and Ordeshook, P. (1970). An Expository Development of a Mathematical Model of the Electoral Process. *American Political Science Review*, 64(2):426.
- Degan, A. and Merlo, A. (2009). Do voters vote ideologically? *Journal of Economic Theory*, 144(5):1868–1894.
- Dow, J. K. and Endersby, J. W. (2004). Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies*, 23(1):107–122.
- Downs, A. (1957). *An Economic Theory of Democracy*. Harper and Row, New York.
- Eguia, J. X. (2012). A spatial theory of party formation. *Economic Theory*, 49(3):549–570.
- Eguia, J. X. (2013). Challenges to the Standard Euclidean Spatial Model. In *Advances in Political Economy*, pages 169–180. Springer, Berlin, Heidelberg, New York.
- Enelow, J. M. and Hinich, M. J. (1984). *The Spatial Theory of Voting: An Introduction*. Cambridge University Press, Cambridge.
- Feddersen, T. (1992). A Voting Model Implying Duverger's Law and Positive Turnout. *American Journal of Political Science*, 36(4):938–962.
- Gabel, M. J. and Huber, J. D. (2000). Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data. *American Journal of Political Science*, 44(1):94–103.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press, Boca Raton.
- Grand, P. and Tiemann, G. (2013). Projection effects and specification bias in spatial models of European Parliament elections. *European Union Politics*, 14(4):497–521.
- Grynaviski, J. D. and Corrigan, B. E. (2006). Specification Issues in Proximity Models of Candidate Evaluation (with Issue Importance). *Political Analysis*, 14(4):393–420.

- Hare, C., Armstrong, D. A., Bakker, R., Carroll, R., and Poole, K. T. (2015). Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions. *American Journal of Political Science*, 59(3):759–774.
- Humphreys, M. and Laver, M. (2010). Spatial Models, Cognitive Metrics, and Majority Rule Equilibria. *British Journal of Political Science*, 40(01):11.
- Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. Wiley, New York.
- Jessee, S. A. (2012). *Ideology and Spatial Voting in American Elections*. Cambridge University Press, Cambridge.
- Kramer, G. H. (1977). A dynamical model of political equilibrium. *Journal of Economic Theory*, 16(2):310–334.
- Krosnick, J. A. (2002). The Challenges of Political Psychology: Lessons to Be Learned from Research on Attitude Perceptions. In Kuklinski, J. H., editor, *Thinking about political psychology*, pages 115–152. Cambridge University Press, Cambridge.
- Laver, M. and Hunt, B. (1992). *Policy and Party Competition*. Routledge, London.
- Lo, J., Proksch, S.-O., and Gschwend, T. (2014). A Common Left-Right Scale for Voters and Parties in Europe. *Political Analysis*, 22(2):205–223.
- Lupia, A. (2016). *Uninformed: Why People Seem to Know So Little about Politics and What We Can Do about It*. Oxford University Press, Oxford.
- McKelvey, R. (1976). Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory*, 12(3):472–482.
- Merrill, S. and Adams, J. (2002). Centrifugal Incentives in Multi-Candidate Elections. *Journal of Theoretical Politics*, 14(3):275–300.
- Merrill, S. and Grofman, B. (1999). *A unified theory of voting. Directional and proximity spatial models*. Cambridge University Press, Cambridge.
- Munger, M. C. and Munger, K. M. (2015). *Choosing in Groups*. Cambridge University Press, Cambridge.
- Patty, J. W., Snyder, J. M., and Ting, M. (2009). Two's company, three's an equilibrium: strategic voting and multicandidate elections. *Quarterly Journal of Political Science*, 4(3):251–278.
- Plott, C. R. (1967). A Notion of Equilibrium and Its Possibility Under Majority Rule. *American Economic Review*, 57(4):787–806.
- Poole, K. T. and Rosenthal, H. (1985). A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, 29(2):357–384.
- Przeworski, A. and Teune, H. (1970). *The Logic of Comparative Social Inquiry*. Wiley, New York.
- Rabinowitz, G. and Macdonald, S. E. (1989). A Directional Theory of Issue Voting. *American Political Science Review*, 83(1):93–121.
- Schofield, N. and Sened, I. (2006). *Multiparty Democracy: Elections and Legislative Politics*. Cambridge University Press, Cambridge.
- Shepsle, K. (1972). The Strategy of Ambiguity. *American Political Science Review*, 66(2):555–568.
- Shikano, S. and Behnke, J. (2009). Issuewählen bei der Bundestagswahl 2005: Eine empirische Schätzung der Verlustfunktion der deutschen Wählerschaft. In Gabriel, O. W., Weßels, B., and Falter, J. W., editors, *Wählen und WÄhler*, pages 251–266, Wiesbaden. VS Verlag.

- Singh, S. (2014). Linear and quadratic utility loss functions in voting behavior research. *Journal of Theoretical Politics*, 26(1):35–58.
- Thurner, P. W. (2000). The empirical application of the spatial theory of voting in multi-party systems with random utility models. *Electoral Studies*, 19(4):493–517.
- Tiemann, G. (2003). Das “most different systems design” als Instrument zum Umgang mit multipler Kausalität. In Pickel, S., Pickel, G., Lauth, H.-J., and Jahn, D., editors, *Vergleichende Politikwissenschaftliche Methoden. Neue Entwicklungen und Diskussionen*, pages 265–287, Opladen. Westdeutscher Verlag.
- Tomz, M. and Van Houweling, R. P. (2009). The Electoral Implications of Candidate Ambiguity. *American Political Science Review*, 103(1):83–98.
- Train, K. E. (2009). *Discrete Choice Methods Model with Simulation*. Cambridge University Press, Cambridge.
- van der Eijk, C., van der Brug, W., Kroh, M., and Franklin, M. (2006). Rethinking the dependent variable in voting behavior: On the measurement and analysis of electoral utilities. *Electoral Studies*, 25(3):424–447.
- Wittman, D. (1973). Parties as Utility Maximizers. *The American Political Science Review*, 67(2):490–498.
- Wittman, D. (1977). Candidates with policy preferences: A dynamic model. *Journal of Economic Theory*, 14(1):180–189.
- Ye, M., Li, Q., and Leiker, K. W. (2011). Evaluating Voter–Candidate Proximity in a Non-Euclidean Space. *Journal of Elections, Public Opinion & Parties*, 21(4):497–521.



Die programmatischen Positionen der deutschen Parteien zur Bundestagswahl 2017: Ergebnisse einer Expertbefragung

Thomas Bräuninger, Marc Debus, Jochen Müller und
Christian Stecker

Zusammenfassung

Die Kurzanalyse berichtet die Ergebnisse einer Expertenumfrage zur programmatischen Ausrichtung der deutschen Parteien im Herbst 2017. Online befragt wurden 93 Politikwissenschaftlerinnen und Politikwissenschaftler. Erhoben wurden die Positionen von acht Parteien auf sechs sachpolitischen Dimensionen sowie die Wichtigkeit dieser Themenbereiche für die Parteien. Im vorliegenden

Danksagung

Wir danken Christian Melbeck für wertvolle Hilfen bei der Programmierung des Expertensurveys.

T. Bräuninger (✉) · M. Debus · C. Stecker
Mannheimer Zentrum für Europäische Sozialforschung (MZES), Mannheim,
Deutschland
E-Mail: thomas.braeuninger@uni-mannheim.de

M. Debus
E-Mail: marc.debus@uni-mannheim.de

C. Stecker
E-Mail: christian.stecker@mzes.uni-mannheim.de

J. Müller
Institut für Politik- und Kommunikationswissenschaft, Greifswald, Deutschland
E-Mail: jochen.mueller@uni-greifswald.de

Beitrag prüfen wir zunächst die Reliabilität des Messinstruments und vergleichen dann die Wichtigkeit der Themenbereiche. Neben der Wirtschaftspolitik wird insbesondere die Zuwanderungspolitik als besonders wichtig eingestuft. Hingegen verlieren, im Vergleich zu früheren Umfragen, die Gesellschaftspolitik und die Umweltpolitik an Bedeutung. Die Positionen der etablierten Parteien sind konsistent mit der Erwartung eines mehrdimensionalen Politikraums, in welchem die Freien Demokraten (FDP) eine wirtschaftsliberale (rechte) Position in der Wirtschafts- und Umweltpolitik und eine progressive (linke) Position in der Gesellschaftspolitik einnehmen. Die Daten erlauben auch eine Einordnung der im 19. Bundestag erstmals vertretenen Alternative für Deutschland (AfD) nach deren Spaltung im Frühsommer 2015. Die AfD wird von den Experten auf allen Dimensionen als rechtskonservative Partei wahrgenommen wird. Fragebogen und Daten werden online zur Verfügung gestellt (Harvard Dataverse, <https://doi.org/10.7910/DVN/22BGVU>), Durchschnittswerte der ermittelten Parteipositionen und parteispezifischen Salienzen der berücksichtigten Politikdimensionen finden sich im Anhang zu diesem Beitrag.

1 Einleitung

Zahlreiche theoretische Ansätze in der Politikwissenschaft nutzen zur Erklärung politischer Prozesse das programmatiche Profil von parteidpolitischen Akteuren – in Form der Position, die die Akteure zu einem Streithema einnehmen, sowie der Bedeutung, die diese dem Thema zumessen (Hinich und Munger 1997). Beispiele sind das auf Downs (1957) zurückgehende räumliche Modell des Parteienwettbewerbs und die zahlreichen darauf basierenden Anwendungen auf das Wählerverhalten und die Determinanten der Positionsverschiebungen von Parteien (etwa Budge et al. 1987, Adams et al. 2005, 2006), die Analyse legislativer Aktivität (etwa Poole und Rosenthal 1997; Hix und Noury 2016) oder von Regierungs- und Koalitionsbildung (Laver und Schofield 1998; Warwick 2006; Pappi 2009). Adäquates Datenmaterial über die Positionen und Salienzzuschreibung einer Partei zu einem bestimmten Streithema ist zentral für die Anwendung der theoretischen Modelle und die Überprüfung der daraus abgeleiteten Hypothesen (Laver 2001). Die analytische Trennung zwischen Positionen und Salienzen ist beispielsweise für die Daten des Manifesto Projects

(MARPOR; vgl. Volkens et al. 2013) nicht gegeben, da die Positionen der Parteien aus deren Themensalienzen abgleitet werden.¹

Angesichts dieses Bedarfs an Daten insbesondere in Folge der Änderung des deutschen Parteiensystems aufgrund der parlamentarischen Präsenz der AfD in Bundestag und Landtagen präsentiert unsere Kurzanalyse die Ergebnisse einer Expertenbefragung, die im November und Dezember 2017 unter Politikwissenschaftlern an deutschen Universitäten und Forschungseinrichtungen durchgeführt wurde. Ziel war es, die partiespezifische Bedeutung von ausgewählten Politikfeldern und die Position der Parteien auf diesen Politikfeldern zu erheben und damit eine empirische Lücke zu füllen. Die letzte Expertenumfrage von vergleichbarem Umfang wurde 2002 von Benoit und Laver (2006) durchgeführt, welche wiederum auf der 1989 durchgeföhrten Untersuchung von Laver und Hunt (1992) aufbaute und diese weiterschrieb. Der Chapel Hill Expert Survey (Polk et al. 2017) hat zwar jüngst neue Daten veröffentlicht, diese werden jedoch nicht im zeitlichen Umfeld von nationalen Wahlen erhoben, was das Heranziehen dieser Daten etwa für die vergleichende Analyse von Regierungsbildung und legislativer Aktivität problematisch machen kann. Die hier präsentierte Umfrage ermöglicht es, den Wandel des deutschen Parteiensystems in den vergangenen knapp 30 Jahren nachzuzeichnen und zu analysieren. Beispielsweise haben sowohl eine veränderte Wichtigkeit sachpolitischer Themen als auch Positionsverschiebungen auf zentralen Dimensionen des politischen Wettbewerbs zu Gründung und Erfolg der Parteien Die Linke und Alternative für Deutschland (AfD) beigetragen. Die Linke profitierte davon, dass die Sozialdemokraten mit der Agenda 2010 des zweiten Kabinetts Schröder (2002–2005) ureigene linke Positionen preisgaben und die Sozial- und Arbeitspolitik stärker in den Fokus der Wähler rückte. Der Erfolg der AfD ist wiederum ohne die Krise der Europäischen Währungsunion und insbesondere ohne die globale Migrationskrise seit 2015 schwer vorstellbar. Unsere Daten erleichtern es, solche langfristigen Entwicklungen nachzuzeichnen, da sich die in der Befragung berücksichtigten Politikfelder an zurückliegenden Surveys orientieren: die bereits erwähnten

¹Es existieren allerdings zahlreiche theoretisch fundierte methodische Ansätze, um von den Salienzen getrennte Positionen aus der Datenstruktur des MARPOR-Datensatzes zu extrahieren (vgl. etwa Pappi und Shikano 2004, 2005; Linhart und Shikano 2007, 2009). Dies ist jedoch – aufgrund der Datenstruktur des MARPOR-Datensatzes – nur für eine begrenzte Anzahl von zwei (Wirtschafts- und Gesellschaftspolitik) oder drei Makropolitikdimensionen (Wirtschafts-, Gesellschafts- sowie Außen- und Sicherheitspolitik) möglich (Debus 2005), sodass sich die Ermittlung politikfeldspezifischer Positionen von Parteien wie etwa zur Bildungs- oder Europapolitik unter Herausrechnung der Salienzen auf der Grundlage des MARPOR-Datensatzes schwierig gestaltet.

Experten-Surveys von Laver und Hunt (1992) und von Benoit und Laver (2006), welche 1989 bzw. 2002 im Feld waren. Dieses Vorgehen maximiert die Vergleichbarkeit der erhobenen Daten zu den Salienzen und Positionen, die von den deutschen Parteien in diesen Politikbereichen jeweils im Kontext von zeitlich nahen Bundestagswahlen angenommen wurden.

Im Folgenden berichten wir zunächst detaillierter über die Durchführung der Expertenbefragung und legen dar, auf welche Politikfelder wir uns konzentrieren. In Abschn. 3 präsentieren wir Evidenz für die Zuverlässigkeit der Experteneinschätzung, um anschließend in den Abschn. 4 und 5 die parteispezifische Wichtigkeit der sachpolitischen Themenbereiche und die Positionen der Parteien zu den berücksichtigten Politikfeldern deskriptiv aufzubereiten. Die über die Expertenbefragung ermittelten Parteipositionen und parteispezifischen Salienzen der berücksichtigten Politikdimensionen werden im Datenanhang präsentiert. In Abschn. 6 untersuchen wir abschließend die Dimensionalität des Konflikttraumes. Unsere Ergebnisse zeigen, dass der deutsche Parteienwettbewerb insofern stabil ist, als er durch zwei grundlegende Konfliktdimensionen gekennzeichnet ist: einen eher sozio-ökonomischen und einen eher gesellschaftspolitischen Konflikt. Beide stehen in engem Zusammenhang mit der allgemeinen Links-Rechts-Dimension, wobei die Positionierung der FDP, aber auch von Die Linke und AfD aus dem Muster herausstechen.

2 Onlinebefragung von Experten

Der Expertensurvey wurde in Form einer Online-Befragung von Politikwissenschaftlern Ende November bis Anfang Dezember 2017 mit dem Programm *limesurvey*² durchgeführt. Die Befragten wurden gebeten, für sieben Themenbereiche einzuschätzen, welche Position die einzelnen Parteien zum Zeitpunkt der Bundestagswahl im Oktober 2017 auf endpunktbezeichneten Skalen von 1 bis 20 einnehmen sowie – jeweils nachfolgend – wie wichtig das entsprechende Thema für die Partei ist.³ Im Anschluss wurde die Einordnung der Parteien auf

²Fragebogen und Datensatz sind über das Harvard Dataverse verfügbar: <https://doi.org/10.7910/DVN/22BGVU>.

³Zum Zweck der Vergleichbarkeit mit den Vorgängerstudien von Laver und Hunt (1992) sowie Benoit und Laver (2006) verwenden wir ebenfalls 20 Antwortkategorien; zu den Vor- und Nachteilen der Verwendung von Mittelkategorien bei bipolaren Ratingskalen siehe Menold und Bogner (2015). Selbstverständlich ist nicht auszuschließen, dass die Einschätzung der Policy-Position einer Partei durch die Experten nicht vollständig unabhängig von der parteispezifischen Bedeutung ist, die die Experten dem Politikfeld zumessen (und umgekehrt). Allerdings sind die ermittelten Positions- und Salienzwerte auf jeden Fall

einer allgemeinen Links-Rechts-Dimension abgefragt. Zum Abschluss sollten die Befragten angeben, wie nahe sie den einzelnen Parteien unter Einbeziehung „aller parteipolitischen Aspekte“ stehen.⁴ Bei allen Fragen sollte die Kategorie „keine Angabe“ gewählt werden, wenn den Befragten eine Einschätzung nicht möglich oder sinnvoll erschien.

Folgende Themenbereiche bzw. Dimensionen wurden abgefragt (mit Angabe der Skalenendpunkte in Klammern):

- Wirtschaftspolitik (1=für Steuererhöhungen/mehr öffentliche Leistungen, 20=für Steuersenkungen/weniger öffentliche Leistungen)
- Gesellschaftspolitik (1=für eine liberale Politik in Fragen wie Abtreibung, Homosexualität und Sterbehilfe; 20=gegen eine liberale Politik in diesen Fragen)
- Zuwanderung (1=Zuwanderung erleichtern, 20=Zuwanderung erschweren)
- Umwelt (1=für Umweltschutz, 20=für Wirtschaftswachstum)
- Bildung (1=für Gemeinschaftsschulen, 20=für mehrgliedriges Schulsystem)
- Europäische Integration (1=mehr Kompetenzen für Europäische Union, 20=weniger Kompetenzen für Europäische Union)
- Links-Rechts (1=links, 20=rechts)

Berücksichtigt wurden die im 19. Bundestag vertretenen Parteien – CDU, CSU, SPD, FDP, Bündnis 90/Die Grünen, Die Linke, AfD – sowie die NPD, die auch in den Befragungen von Laver und Hunt (1992) und Benoit und Laver (2006) einbezogen wurde. CDU und CSU wurden getrennt abgefragt, da nach dem unionsinternen Streit über die Flüchtlingsobergrenze und der Präsentation eines eigenen „Bayernplans“ zur Bundestagswahl 2017 insbesondere für die Zuwanderungspolitik davon auszugehen war, dass sich die Positionen der beiden Schwesterparteien erheblich unterscheiden.

Insgesamt wurden 631 in der politikwissenschaftlichen Forschung aktive Personen kontaktiert, deren E-Mail-Adressen den Rundbriefen der Deutschen Vereinigung für Politikwissenschaft der vergangenen Jahre entnommen wurden. Der Zugang zum Survey erfolgte durch einen html-link und ohne Zugangscode. Insgesamt liegen von 93 Befragten vollständige Fragebögen vor. Dazu zählen

unabhängiger voneinander gemessen als bei der Wahlprogrammanalyse von MARPOR, in welcher schlussendlich die Salienztheorie als Grundlage für die Ermittlung der Positionen einer Partei dient (Laver 2001).

⁴Erste explorative Analysen ergeben keine Hinweise auf systematische Antwortverzerrungen durch die Nähe zu Parteien. Eine weitergehende Analyse würde den Umfang dieses Kurzberichts sprengen.

auch Fragebögen, in denen die abschließende Frage nach der persönlichen Nähe zu den Parteien nicht beantwortet wurde. Die durchschnittliche Befragungszeit betrug 10 min. Die Angaben dieser 93 Befragten bilden die realisierte Stichprobe für die hier berichteten Ergebnisse. Weiterhin wurden die Skalenwerte von 14 Befragten zu einzelnen Dimensionen recodiert, da diese die Skalenenden offensichtlich vertauscht hatten.⁵

3 Zuverlässigkeit und Übereinstimmung der Experteneinschätzung

Wir verwenden im Folgenden die Mittelwerte der Angaben der Experten zu den Parteien als Schätzer der latenten Positionen und zugeordneten Salienzen. Zunächst ist zu fragen, ob die Messung der beiden latenten Größen, der Positionen und der Salienzen, zuverlässig ist. Bei der Einschätzung der Experten handelt sich es ja nicht um persönliche Merkmale, Meinungen oder Einstellungen der Befragten, wie bei Fragen nach dem Haushaltseinkommen oder der Parteiidentifikation in allgemeinen Bevölkerungsumfragen. Vielmehr geht es um die mehr oder weniger gute Einschätzung einer latenten, aber eben für alle Befragten identischen Größe, nämlich die Einschätzung der Positionen einer Partei oder der Wichtigkeit eines Themas für eine Partei. Für die Zuverlässigkeit der Messung sind zwei Fragen zu beantworten. Zum einen, ob der Varianz in den Expertenantworten Unterschiede zwischen den Parteien zuordenbar sind, also inwieweit Interrater-Reliabilität vorliegt. Zum anderen, inwiefern die Experten in ihren Urteilen übereinstimmen, also für identische Objekte dieselben Skalenwerte vergeben (Tinsley und Weiss 2000, vgl. auch Lindstädt et al. 2015). Ersteres ist relevant, weil etwa bei den ideologischen Links-Rechts-Positionen eine hohe Übereinstimmung der Ordnung der Parteien für eine hohe Interrater-Reliabilität ausreichend ist, während letzteres beurteilt, ob die Experten die Skala in gleicher Weise benutzen.

Als Maß der Intercoder-Reliabilität verwenden wir den Intraklassenkorrelationskoeffizienten für einzelne fixe Coder.⁶ Im Ergebnis sind die Intraklassenkorrelationen für alle Dimensionen im Bereich 0,71 bis 0,73 für die Positionsskalen und 0,70 bis

⁵Wir gehen von einer Vertauschung der Skalenendpunkte aus, wenn ein Befragter auf einer einzelnen Dimension den Parteien Positionen in umgekehrte Reihung zuordnet als die (übergroße) Mehrheit der Befragten, bei allen anderen Dimensionen jedoch eine zur Mehrheit konsistente Ordnung angeben kann.

⁶Dieser ergibt sich zu $ICC(3, 1) = \frac{BMS - EMS}{BMS + (k-1)EMS}$, wobei BMS und EMS für die Varianz zwischen den Objekten und die Restvarianz stehen (Shrout und Fleiss 1979).

0,72 für die Salienzskalen und liegen damit sämtlich in einem Bereich, in dem von einer guten Beurteilerübereinstimmung (*rater consistency*) ausgegangen wird. Dies zeigt sich auch im Vergleich zu anderen Expertenumfragen zu Parteien. Lindstädt et al. (2015, S. 20) berichten durchschnittliche Werte von 0,6 für die Positions einschätzungen deutscher Parteien in der 2010er Welle des Chapel Hill Expert Surveys (Bakker et al. 2015). Die Reliabilität der Positionsskalen für das deutsche Parteiensystem bei Benoit und Laver (2006) erweist sich mit Korrelationskoeffizienten von 0,74 bis 0,78 ebenfalls als gut. Insgesamt können wir damit von einer guten Reliabilität unseres Messinstruments ausgehen.

Als Maß für die (absolute) Übereinstimmung der Antworten der Experten folgen wir Lindstädt et al. (2015) und berichten *agreement scores* (Finn 1970), welche die beobachtete Varianz in der Einschätzung zur erwarteten Varianz bei Abwesenheit von Expertenübereinstimmung ins Verhältnis setzt.⁷ Abb. 1 zeigt *agreement scores* für die Experteneinschätzungen zu den Surveyfragen für Positionen und Salienzen und zwar einmal hinsichtlich der Einordnung von Parteien, zum anderen hinsichtlich der Themenbereiche. Insgesamt kann von einer guten bis sehr guten Übereinstimmung ausgegangen werden, wobei zwei Muster augenfällig sind. Erstens herrscht eine größere Übereinstimmung hinsichtlich der Positionen; alle Werte liegen oberhalb des Bereichs von 0,5 bis 0,7, in welchem von einer guten Überstimmung ausgegangen werden kann. Weitgehend einig sind sich die Experten offenbar in Einschätzung der Positionen der etablierten Parteien über alle Dimensionen hinweg. Bei NPD, FDP und AfD zeigen die vergleichsweise großen Interquartilabstände an, dass sich die Experten bei einzelnen Themenbereichen weniger einig darüber sind, welche Positionen diese Parteien hier einnehmen. Auf Ebene der Dimensionen gibt es die geringste Übereinstimmung hinsichtlich der europapolitischen Dimension, sehr einig sind sich die Experten jedoch darüber, welche Positionen die Parteien auf der ideologischen Links-Rechts-Dimension einnehmen. Zweitens ergeben sich geringere *agreement scores* für die Einschätzung der Wichtigkeit der Themen. Für einzelne Parteien (Linke, NPD, FDP und AfD) sowie alle Politikdimensionen liegen diese unterhalb des Schwellenwerts von 0,5. Dies bestärkt das Ergebnis der Reliabilitätsanalyse von oben, dass die Beantwortung der Frage nach der Themenwichtigkeit den Experten offenbar schwerer fällt. Lediglich beim Thema der Zuwanderung scheint die (aktuell hohe) Bedeutung des Themas für die Parteien offensichtlich. Zu berücksichtigen ist hier,

⁷Wir bestimmen den *agreement score* zu $r = 1 - \frac{S}{(20^2 - 1)/12}$, wobei S die Varianz in der Einschätzung und der Nenner der Varianz der Gleichverteilung über die 20 Antwortkategorien unserer Skala entspricht.

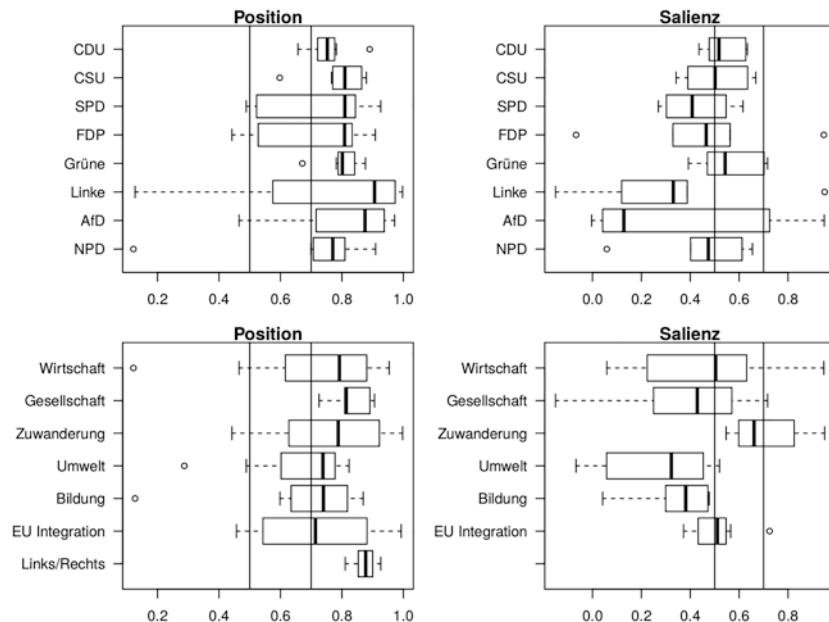


Abb. 1 Übereinstimmung der Experten. Die Abbildung zeigt *agreement scores* der Übereinstimmung der Expertenantworten zu den Surveyfragen für Positionen (linke Spalte) bzw. Salienzen (rechte Spalte) hinsichtlich der Einordnung von Parteien (erste Zeile) und Themenbereichen (zweite Zeile). Dargestellt sind Box-Whisker-Plots (mit Median, Interquartilabstand IQR und $1,5 \times$ IQR sowie Ausreißer)

dass Salienz eine relative Größe ist – die Wichtigkeit eines Themas ergibt sich erst im Vergleich zur Wichtigkeit anderer Themen –, sodass die Unterschiede in den Einschätzungen von einer erwartbaren unterschiedlichen Skalenbenutzung (*differential item functioning*) herrühren können.

4 Wichtigkeit der sachpolitischen Themenbereiche

Welche Themen sind aus Sicht der Parteien besonders wichtig, welche bestimmen den Parteienwettbewerb im Jahr 2017? Wir vergleichen dazu die Salienz der Themen im Durchschnitt über die Parteien, wobei wir der Größe der Parteien Rechnung tragen und die folgenden Themensalienzen als stimmengewichtete Durchschnitte berechnen.

Wie aus Abb. 2 ersichtlich ist, werden die beiden Politikbereiche Wirtschafts- und Sozialpolitik sowie Immigration als die insgesamt für die Parteien wichtigsten Themenbereiche im Herbst 2017 eingeschätzt. An dritter Stelle steht die Europapolitik, gefolgt von der Gesellschafts- und der Bildungspolitik. Der Umweltpolitik kommt mit einem Wert von 11 auf der Skala von 1 bis 20 eine nur mittlere Bedeutung zu. Damit setzt sich ein Trend fort, der sich bereits im Vergleich der Umfragen von 2002 (Benoit und Laver 2006) und 1989 (Laver und Hunt 1992) ergeben hatte: die Gesellschafts- und die Umweltpolitik verlieren an Bedeutung. Wie die Analyse der Positionen der Parteien zeigen wird, liegt dies aber weniger daran, dass mit dem Eingang des Umweltschutzes in das Portfolio aller Parteien und der zunehmenden gesellschaftlichen Akzeptanz von gleichgeschlechtlichen

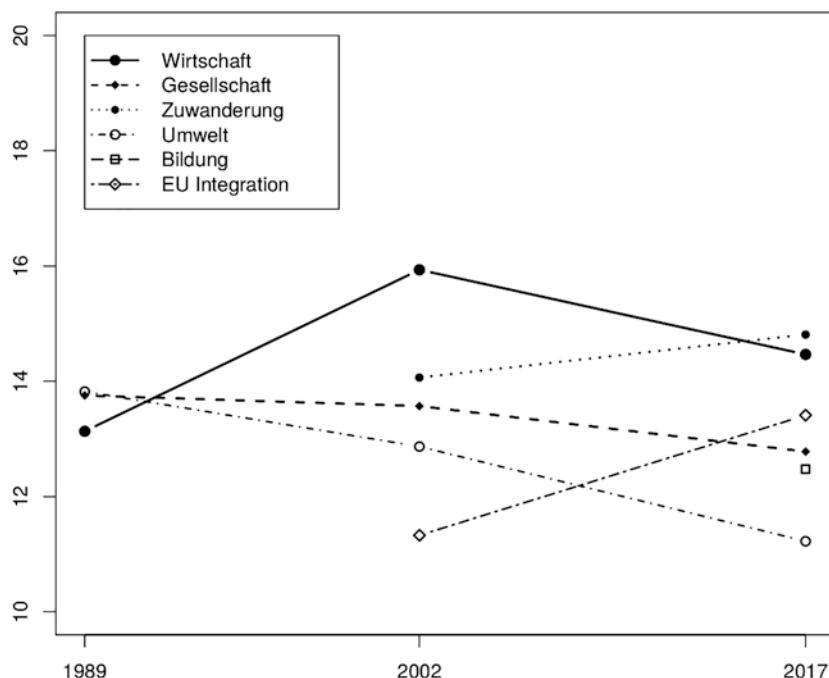


Abb. 2 Wichtigkeit der sachpolitischen Themen im Zeitverlauf. Die Abbildung zeigt die von den Experten wahrgenommene Wichtigkeit der sachpolitischen Themen für die Parteien im Vergleich zu den Expertenumfragen von Laver und Hunt (1992) sowie von Benoit und Laver (2006). Dargestellte Werte sind mit Stimmenanteilen gewichtete, durchschnittliche Salienzwerte auf einer Skala von 1 (unwichtig) bis 20 (sehr wichtig)

Partnerschaften diese Themen konsensfähig wären. Vielmehr scheinen in der Wahrnehmung der Befragten die Parteien wenig Interesse zu haben, diese Themen hervorzuheben. Andererseits stellt, wenig erstaunlich, die Zuwanderungspolitik das wichtigste Politikfeld 2017 dar, nachdem es bereits 2002 als zweitwichtigstes Themenfeld eingeschätzt wurde. Insgesamt eröffnet dies die Frage, ob der für das Parteiensystem der Bundesrepublik lange Zeit charakteristische zweidimensionale Konflikttraum über die Wirtschafts- und Sozialpolitik einerseits und die Gesellschaftspolitik andererseits (Lipset und Rokkan 1967; Pappi 1973, 1977, 1984) hinfällig ist und unterschiedliche gesellschaftspolitische Fragestellungen durch gegensätzliche Auffassungen in Fragen von Migrations- und Integrationspolitik ersetzt werden sollten. Letzteres käme dem Vorschlag von Kriesi et al. (2006) nahe, die von einer neuen Konfliktlinie zwischen Gewinnern und Verlieren von Globalisierungsprozessen ausgehen, womit der Erfolg von links- wie rechts-populistischen Parteien in Europa teilweise erklärt werden könnte.

Wie sich die zugeschriebene Salienz der Themenbereiche zwischen den Parteien unterscheidet, zeigt Abb. 3 (die genauen Werte finden sich in Tab. A1 im Anhang). Drei Beobachtungen sind augenfällig. Erstens unterscheiden sich die

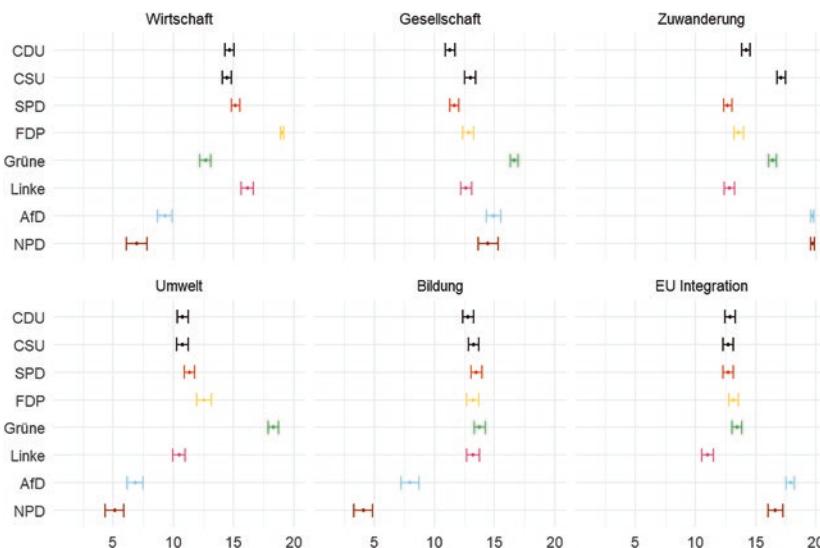


Abb. 3 Wichtigkeit von sechs Themenbereichen für die Parteien im Herbst 2017. Einschätzung durch die Befragten (Mittelwerte und Standardfehler)

Bedeutungsmuster der etablierten Parteien deutlich von denen der AfD und NPD. Bei den etablierten Parteien gibt es wenig Varianz in der zugeschriebenen Bedeutung mit sehr ähnlichen Werten bei Bildung und Europa und der erwartbaren höheren Wichtigkeit der Wirtschaftspolitik für die FDP und der Umweltpolitik für die Grünen. Zweitens werden die Gesellschafts- und insbesondere die Zuwanderungspolitik als zwei Politikfelder gesehen, die für die CSU wichtiger sind als für deren Schwesterpartei CDU. Drittens wird AfD und NPD zugeschrieben, vor allem mit Zuwanderungspolitik und, etwas schwächer, auch mit der Europapolitik befasst zu sein. In der Sicht der Befragten scheinen die Themen Wirtschaft, Umwelt und Bildung eine für AfD wie insbesondere NPD, im Vergleich zu den anderen Parteien, untergeordnete Rolle zu spielen. Auch dies entspricht der allgemeinen Einschätzung der AfD als Ein-Themen-Partei mit der Wandlung von einer euro- zu einer zuwanderungskritischen Partei.

5 Positionierung der Parteien auf den sachpolitischen Dimensionen

Im Vergleich zu den Salienzen sehen die Befragten bei den Positionen deutlichere Unterschiede zwischen den Parteien (Abb. 4, siehe auch Tab. A2 im Anhang). Hinsichtlich aller sechs sachpolitischen Themenbereiche und auch hinsichtlich der ideologischen Links-Rechts-Dimension sehen die Befragten eine Polarisierung mit zumindest einer Partei, die linke oder rechte extreme Standpunkte vertritt. Auch wenn nicht gänzlich auszuschließen ist, dass die Befragten den Skalenraum einfach ausschöpfen (was wiederum die Vergleichbarkeit der Angaben über die Zeit erschweren würde), spricht doch die Verwendung der endpunktbezeichneten Skala gegen eine idiosynkratische Interpretation der Skalen. Insgesamt weisen die Ergebnisse sowohl auf tradierte und bekannte als auch auf neue Konfliktkonstellationen hin.

In der Wirtschaftspolitik stehen sich – wie zu erwarten – ein linkes und ein rechtes wirtschaftspolitisches Lager gegenüber. Die linke Extremposition nimmt Die Linke ein, gefolgt von SPD und Grünen, rechts der Mitte werden die beiden Unionsparteien verortet, die FDP findet sich nahe am rechten Extremstandpunkt, der mit Steuersenkungen und weniger öffentlichen Leistungen einhergeht. Interessant ist hier die Position der AfD, die mit ihren Forderungen nach einer Abgabenbremse für Steuern einerseits, aber auch einer Verlängerung der Bezugsdauer des Arbeitslosengeldes keine eindeutig wirtschaftsliberale Position vertritt, was dem programmatischen Profil rechtspopulistischer Parteien wie etwa der Freiheitlichen Partei Österreichs (FPÖ) oder dem französischen Front National

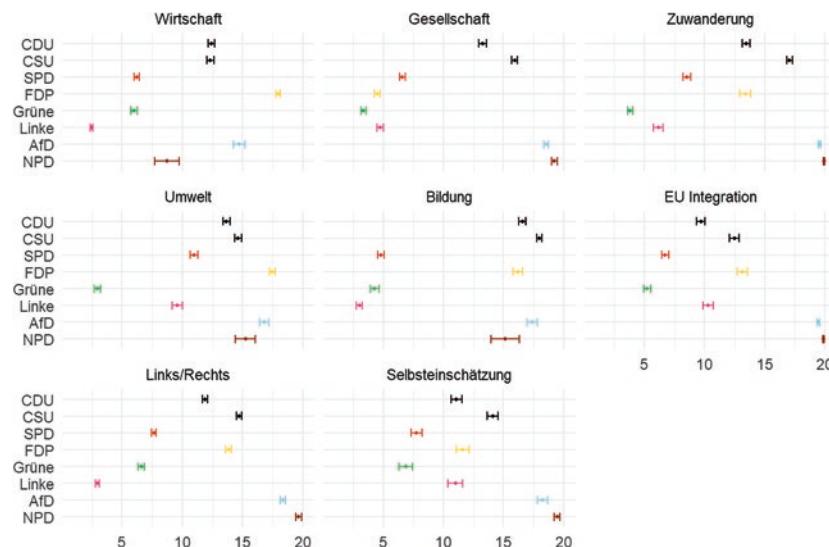


Abb. 4 Positionen der Parteien auf sechs sachpolitischen Dimensionen im Herbst 2017. Einschätzung durch die Befragten (Mittelwerte und Standardfehler)

(FN) entspricht (Decker 2018, S. 359–362). Die NPD wiederum wird mit ihrer gleichzeitigen Ablehnung von „Staatskapitalismus“ und „Liberalkapitalismus“ eher links der Mitte wahrgenommen.⁸

Auch in der Gesellschaftspolitik ergibt sich für die etablierten Parteien der bekannte Konflikttraum. Neben Linken, Grünen und SPD wird auch die FDP als Befürworterin einer liberalen Politik in Fragen wie Abtreibung, Homosexualität und Sterbehilfe angesehen. Demgegenüber lehnen AfD und NPD dies ab. Deutliche Unterschiede gibt es zwischen den Unionsparteien: der CDU werden hier liberalere Positionen zugeordnet als ihrer bayerischen Schwesterpartei. Ähnlich

⁸Die ermittelten Einstufungen von AfD und NPD auf der hier spezifizierten, sowohl Steuer- und Finanzpolitik als auch Wohlfahrtsstaatspolitik kombinierenden sozioökonomischen Dimension legen nahe, dass in künftigen Expertenbefragungen zwischen diesen beiden Politikfeldern unterschieden werden sollte. Natürlich muss bei der Konzipierung von Befragungen auch bedacht werden, dass eine Vergrößerung der abgefragten Politikfelder die Gefahr birgt, dass die Bereitschaft der Teilnahme an der Befragung sinkt.

verhält sich die Konfliktstruktur im aktuell wichtigen Thema der Zuwanderung mit der Konfrontation von linkem und rechtem Lager und Spaltung der Union. Allerdings gibt es eine Ausnahme: Der FDP wird hier eine rechte Position zugeordnet. Angesichts ihrer Forderung nach einem an volkswirtschaftlichen Bedürfnissen ausgerichteten Einwanderungsgesetz erscheint dies plausibel und entspricht auch dem programmatischen Profil weiterer westeuropäischer Parteien mit wirtschaftsliberaler Ausrichtung wie etwa der niederländischen Volkspartij voor Vrijheid en Democratie (VVD).

Die umweltpolitische Dimension weist eine der Wirtschaftspolitik ähnliche Konfliktstruktur auf, lediglich die Grünen nehmen eine pointiert linke Position zum Zielkonflikt zwischen mehr Umweltschutz versus mehr Wirtschaftswachstum ein. Für das Politikfeld Bildung liefert unsere Expertenumfrage, unseres Wissens zum ersten Mal, eine solide Datengrundlage zur Verortung der deutschen Parteien. Die von uns formulierte Frage spannt den Regelungsbereich zwischen den beiden Extremen „für Gemeinschaftsschulen“ einerseits und „für ein mehrgliedriges Schulsystem“ andererseits auf, ein zentrales und kontroverses Thema der Bildungspolitik der letzten Jahrzehnte. Im Gegensatz zu allen anderen sachpolitischen Themen besteht hier eine Regelungskompetenz der Länder, deren Schulsysteme deutlich, teils aus historischen Gründen, variieren. Umso bezeichnender ist die von den Befragten wahrgenommene Polarisierung in der Bildungspolitik. Linke, Grüne und SPD werden als Befürworter von Gemeinschaftsschulen betrachtet, FDP, Union, AfD und NPD als Gegner. Trotz der hohen Polarisierung sollte das Konfliktpotenzial nicht überschätzt werden. Wie oben gezeigt, rangiert die Bildungspolitik bei der Wichtigkeit an zweitletzter Stelle. Dabei gilt anzumerken, dass in der Umfrage nach dem programmatischen Profil der Bundesparteien gefragt wurde – und nicht nach dem der Landesparteien, die aufgrund der Zuständigkeit der Länder für dieses Politikfeld Aspekte von Bildung und Wissenschaft in Landtagswahlkämpfen zentral hervorheben (vgl. etwa Lewandowsky 2013).

Bei der Frage nach den Kompetenzen der Europäischen Union (EU) verläuft die augenfälligste Trennlinie zwischen den beiden rechten Parteien einerseits und den etablierten Parteien andererseits. AfD und NPD werden als klare Befürworter eines Abbaus von Kompetenzen der EU gesehen. Die übrigen Parteien finden sich allesamt bei mittigen Positionen. Eher für einen Ausbau von Kompetenzen der EU sind Grüne und SPD. Linke und CDU werden auf der Mitte der Skala verortet, CSU und FDP leicht rechts davon.

Auch hinsichtlich der Positionierung der AfD auf der allgemeinen, ideologischen Links-Rechts-Dimension sind sich die Befragten weitgehend einig. Die AfD wird klar als Partei am rechten Ende des Spektrums wahrgenommen, lediglich die NPD wird weiter rechts verortet. Bei den etablierten Parteien

werden wiederum die beiden Lager deutlich: im linken Spektrum Linke, Grüne und SPD, rechts von der Mitte CDU, FDP und CSU. Bemerkenswert sind zum einen die deutlich unterschiedlichen Positionen der beiden Unionsparteien, zum anderen die Positionierung der Liberalen rechts der CDU. Damit erweist sich die Links-Rechts-Dimension als weitgehend stabil. Im Vergleich zur Positionierung in der Expertenumfrage von 2002 (die allgemeine ideologische Dimension wurde 1989 nicht abgefragt) ergeben sich nahezu identische Positionen, allerdings mit leicht erhöhter Polarisierung. Eine zentrale Ausnahme bildet die FDP, deren Links-Rechts-Position in der 2002 durchgeföhrten Befragung von Benoit und Laver (2006) mit einem Wert von 13,4 noch leicht links von der Position von CDU/CSU (13,6) angesiedelt war. Gemäß den Ergebnissen der hier präsentierten Erhebung vom Herbst 2017 liegen die Freien Demokraten – auch wenn man CDU und CSU als gemeinsamen Akteur betrachtet – nun rechts der Unionsparteien.

6 Dimensionalität des Politikraums und Vergleich zur ideologischen Links-Rechts-Dimension

Die im Survey abgefragten Dimensionen spannen einen mehrdimensionalen Politikraum auf, in welchem sich die Parteien zu den jeweiligen Streitfragen positionieren. Diese Positionen werden, so zeigen die Ergebnisse der Reliabilitätsanalyse, von den Befragten jeweils auch wahrgenommen. Es gibt keine Hinweise darauf, dass sich die Befragten nicht in der Lage sähen, die Parteien bei den einzelnen Streitfragen zu verorten, ihre Urteile stimmen in großem Maße überein. Das ist insofern nicht überraschend, als es sich bei den Themen um „alte“ Konfliktlinien, wie bei der Frage nach Wirtschaftspolitik, oder, wie der Zuwanderung, um vergleichsweise neue, aber aktuelle Streithemen handelt. Die Analyse der Positionen zeigt auch, dass sich die Ordnung der Parteien auf den einzelnen Politikdimensionen durchaus unterscheidet. Gründung und Aufstieg der AfD können ja gerade dadurch erklärt werden, dass sie eine Lücke im programmatischen Angebot füllt, indem sie bei den beiden aktuellen Themen der Europa- und der Zuwanderungspolitik dezidiert andere Positionen vertritt, als ihr ideologischer Nachbar im konservativ-bürgerlichen Lager, die Union.

Gleichwohl ist zu fragen, ob der mehrdimensionale Issue-Raum auch einen mehrdimensionalen Wettbewerbsraum aufspannt oder ob sich einzelne Themen nicht vielmehr als Variation eines allgemeineren, grundlegenden Konflikts verstehen lassen. Für die Bundesrepublik wird, wie oben erwähnt, mit Lipset und Rokkan (1967) von zwei charakteristischen Spannungslinien ausgegangen, die einerseits den sozioökonomischen Konflikt um die Verteilung von Wohlstand und

Tab. 1 Faktorladungen einer explorativen Faktorenanalyse der Positionseinschätzungen der Befragten (Varimax-Rotation)

	Faktor 1	Faktor 2
	Gesellschaftspolitischer Konflikt	Sozio-ökonomischer Konflikt
Wirtschaft	0,30	0,95
Gesellschaft	0,75	0,20
Zuwanderung	0,85	0,41
Umwelt	0,61	0,50
Bildung	0,63	0,63
EU Integration	0,62	0,30

andererseits den gesellschaftspolitischen Konflikt zwischen säkular-liberalen und religiös-konservativen Gesellschaftsvorstellungen abbilden. Kann auch nach dem Eintritt der AfD in den Bundestag von einem zweidimensionalen Konfliktraum des Parteienwettbewerbs ausgegangen werden?

Die einfache Antwort darauf ist: Ja. Zur Untersuchung des KonfliktRaums führen wir eine Faktorenanalyse der durch die befragten Personen wahrgenommenen Parteipositionen auf den sechs sachpolitischen Dimensionen durch. Dabei zeigt die Extraktion der Eigenvektoren – das Eigenvektorkriterium, ein Scree-Test und eine Parallelanalyse –, dass maximal zwei Faktoren betrachtet werden müssen. Sie erklären 73 % der Varianz in den Daten. Die Faktorladungen sind in Tab. 1 aufgeführt. Auf dem ersten Faktor laden das Gesellschafts- und das Zuwanderungsissue mit Werten über 0,75 sehr stark, beide laden hingegen wenig auf dem zweiten extrahierten Faktor. Umgekehrt korreliert das Wirtschaftsissue stark mit dem zweiten Faktor und kaum mit dem ersten. Es ist daher sinnvoll, den ersten Faktor als letztlich gesellschaftspolitisch latente Konfliktdimension zu begreifen, den zweiten Faktor als sozio-ökonomische Konfliktdimension. Die drei weiteren Themen sind sämtlich mit dem gesellschaftspolitischen Faktor verbunden, mit Werten bis 0,63 allerdings nicht besonders stark. Insbesondere die Frage nach der Stärkung der Europäischen Union scheint eher mit dem Konflikt zwischen gesellschaftlich liberalen und konservativen Auffassungen einher zu gehen, nicht jedoch mit einem wirtschaftspolitischen Konflikt.⁹

⁹Die Skalenhomogenität des ersten Faktors liegt bei 0,9 (Crombach's alpha) und kann durch Herausnahme von Issues nicht verbessert werden.

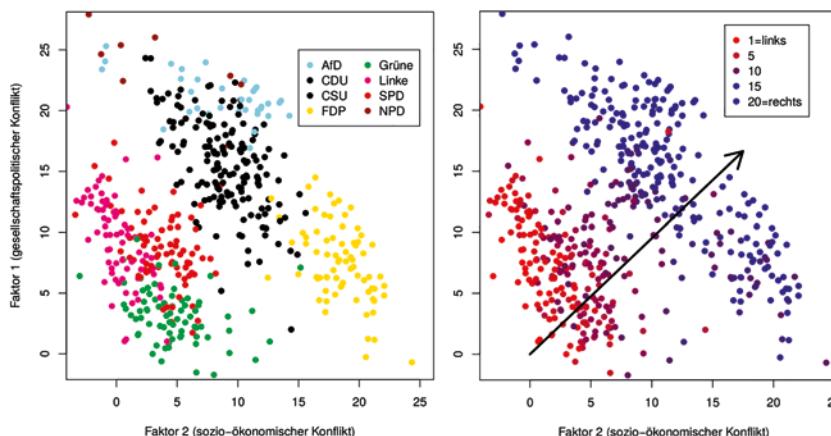


Abb. 5 Wahrnehmene Issue-Positionen der Parteien im zweidimensionalen Faktorenräum (links) sowie deren wahrgenommene Positionen auf der Links-Rechts-Dimension (rechts)

Insofern erweist sich der deutsche Parteidienwettbewerb insofern als stabil, als er auch 2017 durch zwei grundlegende Konfliktdimensionen gekennzeichnet ist. Betrachtet man die wahrgenommenen Parteidpositionen auf den beiden latenten Dimensionen, dann wird der Unterschied zwischen den beiden Konfliktdimensionen besonders deutlich. Das linke Panel von Abb. 5 zeigt die Faktorwerte der durch die Experten wahrgenommenen Positionen der Parteien, also deren Projektion in den durch die Faktorenanalyse erhaltenen zweidimensionalen Politikraum.¹⁰ Auf dem zweiten, sozio-ökonomischen Faktor ergibt sich dieselbe Ordnung der etablierten Parteien von links nach rechts wie beim Wirtschaftssissue: Linke, Grüne, SPD, Unionsparteien und FDP. NPD und AfD werden eher in der Mitte oder links der Mitte verortet, allerdings mit großer Varianz in der Wahrnehmung. Der zweite, gesellschaftspolitische Faktor entspricht in der Ordnung der Parteien dem Gesellschafts- und dem Zuwanderungssissue – Grüne, SPD, Linke und FDP links, Union, AfD und NPD rechts –, wobei sich beide Themen in der Positionierung der FDP unterscheiden.

Der Unterschied wird noch deutlicher, wenn man die Positionierung der Parteien auf der Links-Rechts-Dimension in die Betrachtung einbezieht. Die allgemeine

¹⁰Bezeichnet A die Matrix Faktorladungen und X die Wahrnehmungsdaten, dann ergeben sich die Faktorwerte zu $F = XA'(A'A)^{-1}$.

Links-Rechts-Dimension kann man einerseits als Heuristik verstehen, die wenig informierten Wählern einen Hinweis auf die vermeintliche Positionierung der Parteien zu konkreten Streitthemen gibt. Andererseits ist sie auch *super issue*, welches die Positionen zu einzelnen sachpolitischen Themen zusammenfasst, was wiederum eine Voraussetzung für die Verwendung als Heuristik darstellt. Eine einfache lineare Regressionsanalyse deutet darauf hin, dass die Positionen zur Wirtschafts-, Gesellschafts- und Zuwanderungspolitik besonders wichtig für die Wahrnehmung der ideologischen Positionierung durch die (informierten) Experten sind. Für die drei Politikbereiche ergeben sich (unstandardisierte) Regressionskoeffizienten von 0,44 bzw. 0,21 und 0,24. Demgegenüber sind die Koeffizienten für Umwelt, Bildung und Europa durchgängig kleiner als 0,1. Will man die Links-Rechts-Dimension in Beziehung zum zweidimensionalen Konflikttraum setzen, dann wären die ideologischen Positionen der Parteien entlang der Hauptdiagonalen zu suchen. Im rechten Panel von Abb. 5 sind deshalb die wahrgenommenen Positionen auf der allgemeinen Links-Rechts-Dimension für die einzelnen Stimuli im latenten Konfliktraum abgetragen. Wie erwartet finden sich linke ideologische Positionen vor allem bei latenten Positionen im linken Spektrum des Konflikttraumes, bei Stimuli von Linken und Grünen, sowie moderat linke bei der SPD. Besonders bemerkenswert und hervorzuheben sind hingegen die Verortungen ideologisch rechter Positionen. Extrem rechte Positionen finden sich links oben bei NPD und AfD und teilweise der CSU, gemäßigt rechte Positionen im mittleren Bereich bei CSU und CDU, sowie im rechten unteren Bereich bei Stimuli der FDP. Die Abbildung legt damit drei Befunde nahe: erstens spannt sich, wie erwartet, die ideologische Links-Rechts-Dimension von links unten nach rechts oben im Konflikttraum auf. Zweitens tragen die latenten Konfliktdimensionen, der sozio-ökonomische und gesellschaftspolitische Konflikt, zu gleichen Teilen zur Wahrnehmung der Links-Rechts-Positionierung der Parteien bei. Dies wird nicht zuletzt durch den eingezeichneten Richtungsvektor deutlich, der nahezu parallel zur ersten Hauptdiagonalen verläuft.¹¹ Drittens sind ganz offenbar die „Gründe“, warum eine Partei als ideologisch (moderat) rechts eingestuft wird, vielfältig. Zumindest können sie, wie bei NPD und AfD, mit einer extrem rechten gesellschaftspolitischen Orientierung einhergehen – bei gleichzeitig linken Positionen im sozio-ökonomischen Bereich. Oder, wie bei der FDP, mit wirtschaftlich rechten, marktliberalen Positionen bei gleichzeitig gesellschaftspolitisch liberalen, eben linken Positionen.

¹¹Der Richtungsvektor ergibt sich als Vektor der maximalen Steigung der Abbildung der latenten Positionen auf die ideologischen Positionen.

Eine Beschreibung des Politikraums durch eine einzelne Links-Rechts-Dimension wird den Konfliktmustern des deutschen Parteienwettbewerbs deshalb nicht gerecht. Insbesondere die „Ausnahmepositionierung“ der FDP – als typisch für liberale Parteien in der „Benelux-Konstellation“ (Laver und Hunt 1992, S. 56–57) – wie auch die der AfD werden dabei verdeckt. Eine eindimensionale Perspektive, die auf einem zwischen links und rechts basierenden *super issue* aufbaut, könnte – zum Beispiel bei der Heranziehung der hier präsentierten Daten für handlungs- und entscheidungstheoretische Modelle der Analyse von Wahlverhalten, Parteienwettbewerb, Regierungsbildung, legislativer Aktivität und Politikergebnissen in Form von Policy Outputs und Policy Outcomes – zu verzerrten Ergebnissen führen, die den Einfluss des Policy-Profil der politischen Parteien in Deutschland unter- oder überschätzen. Letzteres könnte dann wiederum zu falschen Schlussfolgerungen für das Ausmaß der demokratietheoretisch hoch relevanten Frage haben, ob die Präferenzen der Bürger – vermittelt in repräsentativen Demokratien über die Parteien und deren programmatisches Profil – sich in den Ergebnissen von Politik widerspiegeln.

Anhang

Tabelle A1 Salienzen der Themenbereiche (Mittelwerte und Standardfehler in Klammern)

	CDU	CSU	SPD	FDP	Grüne	Linke	AfD	NPD
Wirtschaft	14,67 (-0,37)	14,45 (-0,36)	15,17 (-0,37)	19,02 (-0,14)	12,67 (-0,47)	16,16 (-0,5)	9,3 (-0,55)	6,98 (-0,58)
Gesell-schaft	11,3 (-0,41)	12,98 (-0,47)	11,68 (-0,37)	12,83 (-0,46)	16,65 (-0,32)	12,63 (-0,45)	14,92 (-0,56)	14,46 (-0,64)
Zuwan-derung	14,19 (-0,36)	17,11 (-0,34)	12,67 (-0,35)	13,58 (-0,4)	16,39 (-0,33)	12,83 (-0,4)	19,71 (-0,14)	19,72 -0,13
Umwelt	10,77 (-0,45)	10,76 (-0,49)	11,35 (-0,43)	12,53 (-0,62)	18,31 (-0,41)	10,48 (-0,5)	6,84 (-0,6)	5,12 (-0,56)
Bildung	12,8 (-0,43)	13,26 (-0,43)	13,49 (-0,46)	13,2 (-0,49)	13,76 (-0,44)	13,22 (-0,51)	7,95 (-0,59)	4,09 (-0,48)
EU Integ-ration	12,87 (-0,42)	12,71 (-0,41)	12,72 (-0,43)	13,17 (-0,42)	13,44 (-0,39)	10,99 (-0,47)	17,88 (-0,31)	16,63 (-0,47)

Tabelle A2 Parteidpositionen (Mittelwerte und Standardfehler in Klammern)

	CDU	CSU	SPD	FDP	Grüne	Linke	AfD	NPD
Wirtschaft	12,41 (-0,28)	12,32 (-0,29)	6,22 (-0,23)	17,96 (-0,18)	6 (-0,27)	2,46 (-0,13)	14,72 (-0,44)	8,72 (-0,56)
Gesell-schaft	13,25 (-0,31)	15,91 (-0,26)	6,59 (-0,25)	4,52 (-0,26)	3,34 (-0,21)	4,76 (-0,26)	18,54 (-0,18)	19,21 (-0,18)
Zuwan-derung	13,48 (-0,32)	17,08 (-0,22)	8,54 (-0,33)	13,4 (-0,45)	3,84 (-0,21)	6,18 (-0,4)	19,57 (-0,1)	19,94 (-0,03)
Umwelt	13,67 (-0,3)	14,62 (-0,28)	10,99 (-0,32)	17,46 (-0,25)	2,96 (-0,28)	9,58 (-0,43)	16,79 (-0,31)	15,24 (-0,5)
Bildung	16,57 (-0,29)	17,96 (-0,22)	4,79 (-0,27)	16,18 (-0,38)	4,3 (-0,34)	3,02 (-0,24)	17,39 (-0,32)	15,14 (-0,56)
EU Integ-ration	9,7 (-0,35)	12,48 (-0,38)	6,73 (-0,29)	13,14 (-0,44)	5,24 (-0,27)	10,31 (-0,43)	19,48 (-0,1)	19,89 (-0,05)
Links/ Rechts	11,88 (-0,2)	14,74 (-0,21)	7,64 (-0,18)	13,86 (-0,24)	6,58 (-0,26)	2,95 (-0,16)	18,38 (-0,21)	19,67 (-0,22)

Literatur

- Adams, James F., Michael Clark, Lawrence Ezrow und Garrett Glasgow. 2006. Are niche parties fundamentally different from mainstream parties? The causes and the electoral consequences of Western European parties' policy shifts, 1976–1998. *American Journal of Political Science* 50(3): 513–529.
- Adams, James F., Samuel Merrill II und Bernard Grofman. 2005. *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors*. Cambridge: Cambridge University Press.
- Bakker, Ryan, Cathrine De Vries, Erica Edwards, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen und Milada A. Vachudova. 2015. Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010. *Party Politics* 21(1): 143–152.
- Benoit, Kenneth und Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.

- Budge, Ian, David Robertson und Derek Hearl. Hrsg. 1987. *Ideology, strategy and party change: spatial analyses of post-war election programmes in 19 democracies*. Cambridge: Cambridge University Press.
- Debus, Marc. 2005. Annähern oder Distanzieren?: strategische Positionierung von FDP, VdU und FPÖ zwischen 1949 und 2002. *Österreichische Zeitschrift für Politikwissenschaft* 34(4): 413–430.
- Decker, Frank. 2018. Was ist Rechtspopulismus? *Politische Vierteljahresschrift* 59(2): 353–369.
- Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper & Row.
- Finn, Robert A. 1970. A note on estimating the reliability of categorical data. *Educational and Psychological Measurement* 30(1): 71–76.
- Hinich, Melvin J. und Michael J. Munger. 1997. *Analytical politics*. Cambridge: Cambridge University Press.
- Hix, Simon und Abdul Noury. 2016. Government-opposition or left-right? The institutional determinants of voting in legislatures. *Political Science Research and Methods* 4(2): 249–273.
- Kriesi, Hanspeter, Edgar Grande, Romain Lachat, Martin Dolezal, Simon Bornschier und Timotheos Frey. 2006. Globalization and the transformation of the national political space: Six European countries compared. *European Journal of Political Research* 45(6): 921–956.
- Laver, Michael. 2001. Position and salience in the policies of political actors. In *Estimating the policy position of political actors*, Hrsg. Michael Laver, 66–76. London/New York: Routledge.
- Laver, Michael und Ben W. Hunt. 1992. *Policy and party competition*. New York: Routledge.
- Laver, Michael und Norman Schofield. 1998. *Multiparty government: The politics of coalition in Europe*. Ann Arbor: University of Michigan Press.
- Lewandowsky, Marcel. *Landtagswahlkämpfe: Annäherung an eine Kampagnenform*. Wiesbaden: Springer-Verlag, 2013.
- Lindstädt, René, Sven-Oliver Proksch und Jonathan B. Slapin. 2015. *Assessing the measurement of policy positions in expert surveys*. Konferenzpapier, Biennial Conference of the European Union Studies Association, Boston.
- Linhart, Eric und Susumu Shikano. 2007. *Die Generierung von Parteipositionen aus vorverschlüsselten Wahlprogrammen für die Bundesrepublik Deutschland (1949–2002)*. Mannheimer Zentrum für Europäische Sozialforschung, Arbeitspapier Nr. 98.
- Linhart, Eric und Susumu Shikano. 2009. Ideological signals of German parties in a multi-dimensional space: An estimation of party preferences using the CMP data. *German Politics* 18(3): 301–322.
- Lipset, Seymour M. und Stein Rokkan. 1967. Cleavage structures, party systems, and voter alignments: an introduction. In *Party Systems and Voter Alignments: Cross-National Perspectives*, Hrsg. Seymour M. Lipset und Stein Rokkan, 1–64. New York: The Free Press.
- Menold, Natalja und Bogner, Kathrin. 2015. *Gestaltung von Ratingskalen in Fragebögen*. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (SDM Survey Guidelines). https://doi.org/10.15465/sdm-sg_015.

- Polk, Jonathan, Jan Rovny, Ryan Bakker, Erica Edwards, Liesbet Hooghe, Seth Jolly, Jelle Koedam, Filip Kostelka, Gary Marks, Gijs Schumacher, Marco Steenbergen, Milada Vachudova and Marko Zilovic. 2017. Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics* 4(1): 1–9.
- Pappi, Franz U. 1973. Parteiensystem und Sozialstruktur in der Bundesrepublik. *Politische Vierteljahresschrift* 14(2): 191–213.
- Pappi, Franz U. 1977. Sozialstruktur, gesellschaftliche Wertorientierungen und Wahlabsicht: Ergebnisse eines Zeitvergleichs des deutschen Elektorats 1953 und 1976. *Politische Vierteljahresschrift* 18(2–3): 195–229.
- Pappi, Franz U. 1984. The West German party system. *West European Politics* 7(4): 7–26.
- Pappi, Franz U. 2009. Regierungsbildung im deutschen Fünf-Parteiensystem. *Politische Vierteljahresschrift* 50(2): 187–202.
- Pappi, Franz U. und Susumu Shikano. 2004. *Ideologische Signale in den Wahlprogrammen der deutschen Bundestagsparteien 1980 bis 2002*. Mannheimer Zentrum für Europäische Sozialforschung, Arbeitspapier Nr. 76.
- Pappi, Franz U. und Susumu Shikano. 2005: Regierungsabwahl ohne Regierungsneuwahl. *Politische Vierteljahresschrift* 46(4): 513–526.
- Poole, Keith T. und Howard Rosenthal. 1997. *Congress: A political-economic history of roll call voting*. Oxford: Oxford University Press.
- Shrout, Patrick E. und Joseph L. Fleiss. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* 86(2): 420–428.
- Tinsley, Howard E. A. und David J. Weiss. 2000. Interrater reliability and agreement. In *Handbook of applied multivariate statistics and mathematical modeling*, Hrsg. Howard E. A. Tinsley und Steven D. Brown, 95–124. Amsterdam: Elsevier.
- Volkens, Andrea, Judith Bara, Ian Budge, Michael D. McDonald und Hans-Dieter Klingemann. 2013. *Mapping Policy Preferences from Texts: Statistical Solutions for Manifesto Analysts*. Oxford: Oxford University Press.
- Warwick, Paul V. 2006. *Policy horizons and parliamentary government*. Cambridge: Cambridge University Press.

Teil II

Kurzanalysen



Perceived Party Placements and Uncertainty on Immigration in the 2017 German Election

Ingrid Mauerer and Micha Schneider

Abstract

Almost all national election studies contain policy scales that are intended to measure where respondents perceive parties or candidates on central campaign issues. These placements form the basis for models of survey responses, party perceptions, and voter choice. It is well known that the placements might be affected by uncertainty. We use the finite mixture model ‘BetaBin’ to study response patterns to party placements on policy issues. The model consists of a placement part and an uncertainty part. Whereas the placement part of the model accounts for lower and higher placements on the ordinal scales, the uncertainty component accounts for tendencies to locate the parties on the middle or at the extremes of the policy scales. We use the 2017 German national election and apply the model to the immigration issue. Our results demonstrate that uncertainty strongly influences the respondents’ perceptions of most parties. Neglecting this structure leads to worse models as indicated by performance measures.

The order of the authors is alphabetical. Both authors contributed equally.

I. Mauerer (✉)

Geschwister-Scholl-Institut für Politikwissenschaft, Munich, Germany

E-Mail: Ingrid.Mauerer@gsi.uni-muenchen.de

M. Schneider

Institut für Statistik, Munich, Germany

E-Mail: Micha.Schneider@stat.uni-muenchen.de

Keywords

Party Placements · Uncertainty · Mixture Models · BetaBin · 2017 German National Election

1 Introduction

Almost all national election studies around the world contain policy scales. These policy scales were developed to assess the policy preferences of citizens and the positions of political figures on central campaign issues. A lot of theoretical and empirical concepts rely on the respondents' stated policy preferences and political perceptions. They form the basis for models of survey responses, party perceptions, policy choices, and voter choice. In public opinion research, several studies explore citizens' policy preferences (i.e., self-placements) on policy scales and test rival explanations of variability in attitudes due to uncertainty, ambivalence or equivocation. Alvarez and colleagues focus on policy choices of American citizens on public policies such as abortion (Alvarez and Brehm 1995), racial policies (Alvarez and Brehm 1997) or Internal Revenue Services (Alvarez and Brehm 1998; see also Alvarez and Brehm 2002). In a similar vein, Harbers et al. (2013) explore response variability in Left-Right placements among the Latin American electorate, and De Vries and Steenbergen (2013) examine European citizens' ambivalence in attitudes toward European integration.

Citizens' policy preferences and their perceptions about party platforms also play a central role in spatial voter choice models. These models assume that parties take stances on issues and that voters can perceive these stances (Downs 1957; Davis et al. 1970; Campbell et al. 1960). However, Shepsle (1972) and Enelow and Hinich (1981) reasoned that voters might be uncertain about the positions parties or candidates take on policies due to limited information on the side of the voters or position blurring on the side of the parties. They propose to represent party or candidate positions by probability distributions instead of single points. A few empirical studies account for voter uncertainty and incorporate it into the choice rule (e.g., Bartels 1986; Gill 2005; Berinsky and Lewis 2007). These studies highlight that electoral decisions are not only systematically related to spatial proximity but also to uncertainty about party platforms. Another recent study found that also voters are quite uncertain about their policy preferences and show a considerable degree of inconsistency (Stoetzer 2017). In addition, not all parties put the same emphasis on the same issues and the reliance on issues when voting is party-specific (Mauerer et al. 2015; Mauerer 2016). As a result, citizens might be uncertain what position they should ascribe when it comes to specific issues and parties.

If ordinary citizens do elect parties or candidates that best represent their preferences on public policies, it is necessary to understand in the first place how they perceive the parties' policy platforms and what role uncertainty plays in these perceptions. As uncertainty about party platforms can hamper democratic representation, we need models of survey responses that are able to detect the impact of uncertainty and to account for different response patterns due to uncertainty. Such insights will help us to understand how the electorate incorporates uncertainty into their political perceptions. This will then add to our understanding of the impact of policy-oriented decision making and its electoral consequences.

The approach we develop in this paper is a model of political perceptions that allows detecting how the perceptions of party platforms among ordinary citizens are structured. It is a model of survey responses to ordinal policy scales where specific response styles capture the uncertainty structure. The model belongs to the class of mixture models for ordinal data that are able to account for both the placement structure and the uncertainty structure of responses. We apply the so-called BetaBin model (Tutz and Schneider 2018) that can handle different response patterns when citizens are uncertain where to place the parties on the policy scales: a tendency to the middle category and a tendency to extreme categories. Whereas the concentration in the middle category is widely known (see, e.g., Aldrich et al. 1982; Alvarez and Franklin 1994), the approach we develop does not only account for this tendency but also for response styles to extreme categories, which has been discussed by Baumgartner and Steenkamp (2001) for example. Both the placement and the uncertainty structure of responses can be related to covariates. Since there are little theoretical and empirical insights into the mechanisms of how respondents' political perceptions are structured, we exploratively evaluate different sets of explanatory variables. We examine predictors that relate to cognitive processes or the respondents' information costs, to the relationship between the respondent and the party to be located at the policy, to issue characteristics, and standard demographics.

Compared to existing approaches, we see three main advantages of the proposed model. First, instead of modeling the variance by additional scale parameters as done in the heteroscedastic regression model (e.g., Harvey 1976; Alvarez and Brehm 1995), our approach can model specific response styles, namely the tendency to the middle, random choice or to extreme categories. Thus, we are able to detect particular structures of uncertainty which can be explained by covariates. Using models with scale parameters can only model high or low variance, but the variability is still rather unstructured. Second, the model is designed for ordinal responses, as compared to previous work relying on the logit or probit model for binary response (e.g., Alvarez and Brehm 1995) or on the linear regression

model (e.g., Harbers et al. 2013; De Vries and Steenbergen 2013). It is well known that these models are not the best choice for modeling ordinal response data. Third, we do not need additional survey questions that directly ask respondents how certain they are about party or candidate positions, which are very rare in surveys.

We use the 2017 German national election study (Rossteutscher et al. 2017) to demonstrate the advantages of the proposed model. The election study contains typical eleven-point issue scales to measure the positions of parties on issues of current concern, such as immigration, taxes and climate change. On these ordinal scales, respondents were asked to place the parties. We apply the model to the immigration issue that played a significant role in the 2017 election campaign with different parties being more or less clear or ambiguous in the position they offered on it. We examine where the respondents perceive the major German parties on this central campaign issue and what role uncertainty plays in these perceptions. Our results show that the BetaBin model provides fruitful and new insights into the perceptions of party platforms and outperforms traditional cumulative models without uncertainty structure. Including the uncertainty structure leads to better model performances, and therefore increases our understanding of political perceptions. Uncertainty strongly determines the respondents' perceptions of most parties on the immigration issue. Whereas the respondents expressed a clear preference where to place the AfD, they exhibit major difficulties in locating the CDU and the FDP.

2 Measuring and Modeling Uncertainty

There are some empirical approaches in the literature on how to measure and model uncertainty in party platforms. Here, we give a very brief overview of the most important models and approaches. One way to deal with variability and uncertainty in survey responses is to rely on range formats that adjust the traditional seven-point or eleven-point policy scales (see, e.g., Tomz and Van Houweling 2009; Aldrich et al. 1982; Alvarez 1999). Another approach is to stick to the original policy scales and to design additional survey questions to measure and examine uncertainty variability of survey responses (e.g., Alvarez and Franklin 1994). These questions directly ask respondents how certain they are about candidate or party positions after they have provided these placements. However, only a few electoral studies have included self-reports on uncertainty yet.

Instead of relying on survey-based measures of uncertainty using self-assessments of respondents, Bartels (1986) proposes the following two-stage procedure to examine the impact of issue uncertainty on individual voting behavior. First, he develops a model of survey responses to assess the respondents' uncertainty

in party placements. He takes refused answers as an indicator of uncertainty. The basic idea is that respondents who are uncertain are not able to give a placement at all. If a respondent refuses to place a candidate, this is interpreted as uncertainty which can be modeled as a function of observable characteristics of the respondent, the candidate, and the political environment. Based on the estimated probabilities of non-response that should capture the variance of candidate perceptions, in the second stage, he estimates a voter choice model to assess the importance of uncertainty in individual voting decisions. In both stages, he uses a linear probability approach.

Another idea is to model the variance by a heteroscedastic regression model introduced by Harvey (1976). In this case, the variance of the disturbance is modeled by covariates. Alvarez and Brehm (1995) apply a heteroscedastic binary probit model to analyze attitudes toward abortion in the U.S. electorate. Harbers et al. (2013) and De Vries and Steenbergen (2013) use a heteroscedastic linear regression model. These approaches have the disadvantage that they are not designed for small ordinal response scales. Using heteroscedastic linear regression for ordinal responses can lead to several difficulties. The error terms might not be normally distributed, and the linear regression might predict values lower, in between or above the response scale. Furthermore, it is not designed to measure specific response styles such as a tendency to middle or extreme categories.

The model by Rozenas (2013) can be seen as a combination of non-response (Bartels 1986) and variance heterogeneity (Harvey 1976) which leads to a complex model with hyper-parameters and the necessity of choosing appropriate prior distributions. Another approach was developed by Gill (2005) who combines uncertainty with the concept of entropy. The entropy approach is based on an aggregate measure of issue uncertainty that uses information on the survey question, the issue to be evaluated, attributes of the candidates or parties as well as aggregated responses by the whole sample. In contrast to Bartels (1986) who imposes a homogeneous uncertainty threshold across respondents to model uncertainty, the entropy approach is based on an uncertainty term that is still the same across respondents but varies across issues and candidates.

The existing approaches deal very differently with missing values. The crucial question when relying on missing data in the response to measure uncertainty is whether there is a particular mechanism for generating the missing data. Empirical applications based on pure heteroscedastic models (as, e.g., Harvey 1976) do not make use of any missing data and rely only on observed values. Contrarily, Bartels (1986) and Rozenas (2013) argue that missing data in the response is caused by the uncertainty of the respondents and related to covariates. This might be the case but maybe not the only or major process of generating missing values in the response structure. Respondents might have a clear position but do not want

to report it because of social desirability, which is quite probable when it comes to delicate questions or policies. Another reason might be that respondents just skip the question because of time limitations or lack of motivation. In such cases, missing values consist of both uncertain and certain placements. Since we usually do not know the true data generating process of missing data, we prefer to exclude the missing values (including ‘don’t know’ answers) from the analysis instead of assuming that missing data in the response is directly linked to the uncertainty of the respondents.

3 Response Styles and Variability in Uncertainty across Parties

The literature is in agreement that uncertainty is inherently subjective and that particular segments are more certain or uncertain about party placements. Previous research on response patterns mainly suspected that respondents show a tendency to the middle of the scale due to limited information, when they are not politically interested or involved. On the policy scales, the middle categories reflect moderate positions. Whereas the concentration in the middle category is widely known (see, e.g., Aldrich et al. 1982; Alvarez and Franklin 1994), the approach we develop in this paper can account for several kinds of uncertainty—especially the tendency to middle or extreme categories. Particularly the response style to extreme categories, i.e., a tendency to ascribe parties extreme policy stances, seems to be very promising because the response patterns might not only be the result of cognitive processes of citizens but might also stem from how parties behave on issues, what strategies parties pursue on particular issues. Parties might take ambiguous stances, blur their positions that induce uncertainty on the side of the voters where the party actually positions itself. However, parties also often overshoot their positions (Kedar 2005a, b). Therefore, one might also expect that respondents show a tendency to place the parties at the extremes when they are uncertain.¹

In addition, we expect differences in uncertainty patterns across parties. In general, parties often pursue different strategies on different issues, yielding different levels of uncertainty in the position respondents ascribe to parties. We argue

¹Baumgartner and Steenkamp (2001) provide other reasons. For instance, extreme response styles (ERS) can be seen as a ‘reflection of rigidity’. Baumgartner and Steenkamp (2001) and Van Vaerenbergh and Thomas (2013) give an overview of different response styles.

that this variation relates to the underlying party system and resulting dynamics of party competition that reduce or increase uncertainty in perceived party platforms. We apply the model to the issue of immigration. Why should there be different levels of uncertainty in the position respondents ascribe to parties on the immigration issue? We expect that there is systematic variation in uncertainty on the immigration issue due to party family.

Immigration lies at the core of Inglehart's (1997) post-materialist dimension, and therefore represents an important 'new politics' issue. Immigrants issues grew increasingly salient in Western Europe in the last decades and gave rise to the emergence of new competitors on the radical right of the political spectrum. A considerable amount of scholarly attention has been paid to the explanation of the electoral fortunes of populist radical right parties (see, e.g., Givens 2005; Kitschelt and McGann 1995; Mudde 2007; Ignazi 2003; Art 2011). Also, much scholarly work has been devoted to studying of the dynamics of party competition on immigrants issues (see, e.g., Meguid 2005; Ivarsaten 2008; Norris 2005; Bale 2003, 2008; Abou-Chadi 2015; Pardos-Prado et al. 2014). The recent studies point out that right-of-center and populist radical right parties have a strategic advantage over their competitors on immigrants issues. These parties lay particular emphasis on their core issue immigration by profoundly polarizing on them. In the light of their issue portfolio and long-term ideological backgrounds, center-right and radical right parties increase the saliency of immigrants issues by strongly politicizing it and taking unambiguous restrictive stances. As a result, one might expect that citizens are quite certain in their perceptions about what positions radical right and right-of-center parties offer on immigration.

Other studies focus on the mainstream left and investigate the electoral strategies of this party family concerning immigration (see, e.g., Bale et al. 2010; Alonso and Da Fonseca 2012). For instance, Bale et al. (2010) examine the strategic responses of social-democratic parties. They find that these parties face a 'triple challenge' due to the rise and the success of extreme right parties: 1) Populist radical right parties mainly campaign on immigrants issues—and therefore increase the saliency of issues—that are traditionally owned by right-of-center parties; 2) the extreme right attempts to mobilize the working-class who is habitually linked to left-of-center parties; and 3) populist radical right parties ease to form center-right governments (see also Bale 2003, 2008). Based on this line of reasoning, one might expect that citizens have difficulties in locating the mainstream left on immigrants issues.

4 A Model of Perceived Party Platforms under Uncertainty

Our modeling approach is based on the idea that variability in survey responses can be modeled by mixture models. In the framework of mixture models, any density f can be represented as a combination of a finite set of densities so that

$$f = \sum_{g=1}^M \pi_g f_g, \quad (1)$$

where $0 \leq \pi_g \leq 1$ is the mixture proportion or weight, and M is the number of densities used to describe the density f . Mixture models are widely used. An introduction to this model class is given by McLachlan and Peel (2000). Iannario and Piccolo (2016a, b) provide an overview of mixture models for ordinal data.

When studying perceived party placements, several requirements have to be considered. First, the number of densities M can be restricted to two, as we are only interested in the placement structure and uncertainty structure of survey responses. Second, we have ordinal responses so that we need density functions that are appropriate for this data type; any continuous densities cannot be considered. Third, we would like to use densities that are the best choice for modeling both components. However, the best choice is not always to use densities from the same type for both components.² By including an uncertainty component, we can account for specific response styles. One basic response style can be represented by a uniform distribution corresponding with a random choice of response category as done for example by D'Elia and Piccolo (2005) or Tutz et al. (2017). We use the so-called BetaBin model proposed by Tutz and Schneider (2018) which is characterized by a flexible modeling of the response style and placement structure.

4.1 Model Formulation

Let R_i be the observed response of an individual i to an ordinal policy scale taking the values $\{1, \dots, k\}$. Y_i denotes an unobserved random variable that presents the deliberate choice, i.e., the real party placement. Let U_i be the unobserved uncertainty component modeling the type of response style. Both Y_i and U_i

²Choosing two binomial distributions are considered as the same type.

take ordered values $\{1, \dots, k\}$. Given explanatory variables \mathbf{x}_i and \mathbf{w}_i , the mixture model ‘BetaBin’ has the form

$$P(R_i = r | \mathbf{x}_i, \mathbf{w}_i) = \pi_i P_M(Y_i = r | \mathbf{x}_i) + (1 - \pi_i) P_U(U_i = r | \mathbf{w}_i). \quad (2)$$

\mathbf{x}_i and \mathbf{w}_i are vectors of covariates for both components that can be identical, overlapping or completely different. π_i represents the mixture probability that measures the importance of the structured component in the mixture model. Thus, the observed response results from a discrete mixture of the placement and the uncertainty component.

For the placement component $P_M(Y_i = r | \mathbf{x}_i)$, any ordinal model would be possible. We use the following widely known cumulative logit model (see Tutz 2012), also known as ordered or ordinal logit model:

$$\log \left(\frac{P(Y_i \leq r | \mathbf{x}_i)}{P(Y_i > r | \mathbf{x}_i)} \right) = \gamma_{0r} + \mathbf{x}_i^T \boldsymbol{\gamma}, \quad r = 1, \dots, k-1,$$

where γ_{0r} are the intercepts or thresholds and $\boldsymbol{\gamma}$ the estimated effects independent of r . Note that in the literature different notations of the cumulative logit model are used. Here a positive value of $\boldsymbol{\gamma}$ indicates that a lower category is more probable. The response style U follows a beta-binomial distribution, $U \sim \text{Beta-binomial}(k | \alpha, \beta)$ with the mass function

$$f(u) = \begin{cases} \binom{k-1}{u-1} \frac{B(\alpha + u - 1, \beta + k - u + 1)}{B(\alpha, \beta)} & u \in \{1, \dots, k\} \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha, \beta > 0$ are the parameters of the beta-binomial distribution. $B(\alpha, \beta)$ is the beta function defined as

$$B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt.$$

Since α and β are not identical with the location and scale of the distribution, the following reformulation is necessary:

$$\mu = \alpha / (\alpha + \beta), \quad \delta = 1 / (\alpha + \beta + 1).$$

Now the expected value $E(U)$ and the variance $\text{var}(U)$ are given by

$$E(U) = (k-1)\mu + 1, \quad \text{var}(U) = (k-1)\mu(1-\mu)[1 + (k-2)\delta].$$

As $\delta \rightarrow 0$, the beta-binomial distribution converges to the (shifted) binomial distribution $B(k, \mu)$ with mean μ and support $\{1, \dots, k\}$. The specific response styles characterized by a tendency to middle or extreme categories are determined by

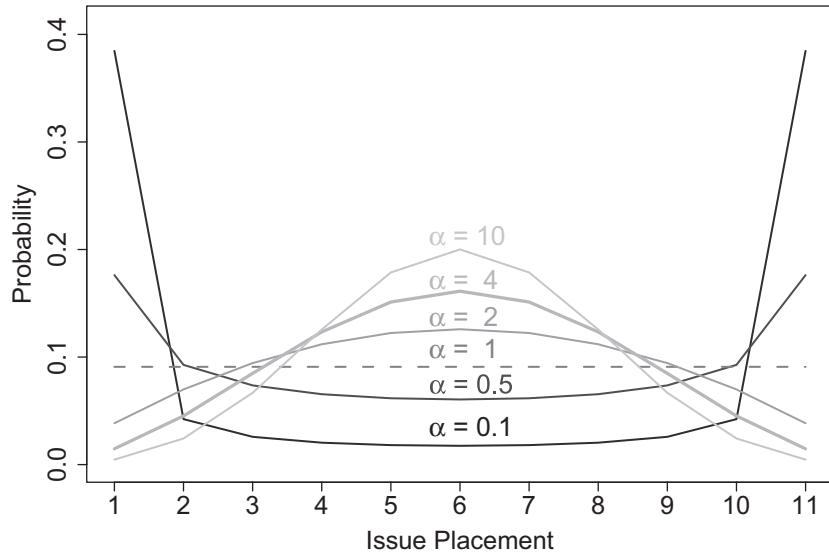


Fig. 1 Probability mass on categories for various values of α for $k = 11$ categories

imposing the restriction $\alpha = \beta$, which leads to $\mu = 0.5$ and $\delta = 1/(2\alpha + 1)$. Thus, the location of the distribution is always fixed at the middle of support. The only flexible parameter is δ or rather α . The smaller α , the larger δ and therefore the variance. Figure 1 shows the different shapes of this restricted beta-binomial distribution.

α values larger than one correspond with a tendency to the middle categories, while α values smaller than one indicate a tendency to extreme categories. The distribution ranges from a (shifted) binomial distribution with the mode in the middle of support to almost a distribution with two equal point mass at the border of support. The first distribution corresponds with a strong tendency to the middle categories, the latter with a strong tendency to both extreme categories given by the minimum and maximum of k . All gradations between the two extreme cases are possible. The parameter α is linked to covariates w_i by

$$\alpha = \exp(w_i^T \alpha) = \exp(\alpha_0) \exp(\alpha_1)^{w_{i1}} \dots \exp(\alpha_m)^{w_{im}},$$

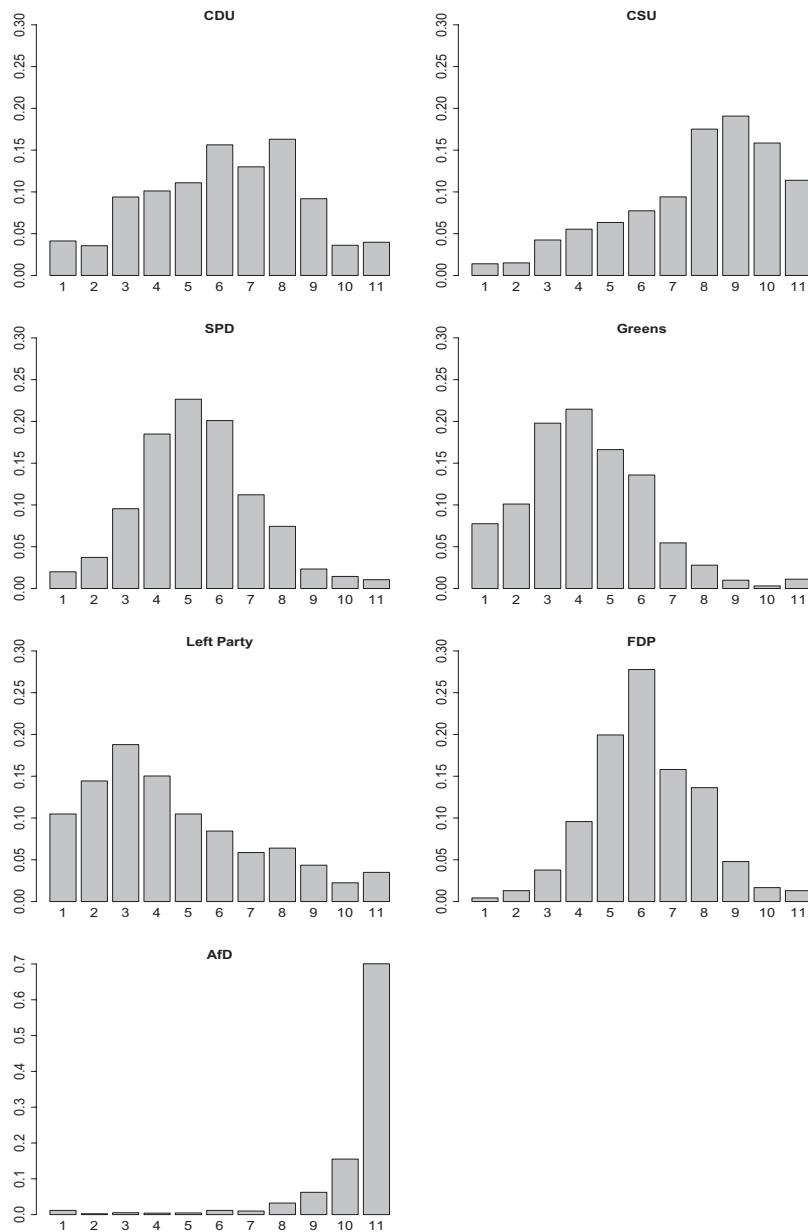
where α is the parameter of the restricted beta-binomial distribution and α_j gives the effect of the j -th covariate linked with $\exp()$ to α . Thus, α changes by the factor $\exp(\alpha_j)$ when w_{ij} increases by one unit, given all other variables in the model

are kept constant. The exponential function ensures that α is always positive, although the effects of the covariates may be positive or negative. Positive α_j values lead to α values larger than one and indicate a tendency to middle categories. Negative α_j values lead to α -values smaller than one and indicate the tendency to extreme categories. For example, an effect of $\alpha_j = 1$ leads to $\alpha = \exp(1) = 2.71$ showing a tendency to the middle categories for the j -th covariate. For $\alpha = 1$ ($\alpha_j = 0$) one obtains the discrete uniform distribution (dashed line in Fig. 1) corresponding with a random choice of a category.

4.2 Survey Responses to Party Placements

We apply the model of party perceptions to the immigration issue contained in the 2017 German national election study (Rossteutscher et al. 2017). The respondents were asked to state where they perceive the parties on an eleven-point scale with the following endpoints: 1 “Immigration should be facilitated” (pro) and 11 “Immigration should be restricted” (contra). We restrict our analysis to the seven most important German parties: the Christian Democratic Union (CDU), its Bavarian sister party the Christian Social Union (CSU), the Social Democratic Party (SPD), the Liberal Party (FDP), the Greens, the Left Party, and the Alternative for Germany (AfD). We also excluded respondents that provided no answer or opted for the ‘don’t know’ category. The stated positions of these parties on the immigration issue present the observed response R_i in Eq. 2.

Figure 2 illustrates the distribution of party perceptions on the immigration issue. Since not all survey respondents were able to locate all the seven parties, the number of observations (out of 2179 total respondents) slightly differs. The minimum number of observations is 1387 for the FDP, and the maximum is 1949 for the CDU. For each of the eleven categories, the percentages are reported in Fig. 2 so that the shape between the different parties is comparable even though the absolute numbers are not identical. We observe that the shape of the distributions is very different between the seven parties. Unsurprisingly, almost 70% of the respondents locate the AfD at the far right, resulting in a very skewed distribution. Also, the CSU is perceived as taking a rather contra-immigration stance, whereas the respondents place its sister party CDU closer to the middle categories without a clear modal value. The distributions of the perceptions for the FDP and SPD are more symmetric with modal values at 6 and 5, respectively. The perceptions of the Greens are skewed toward the pro-immigration pole, and the Left Party is perceived as taking the most pro-immigration stance. This data situation demands a flexible model that can handle all these different distributions, and therefore presents an ideal situation to demonstrate the benefits of the proposed mixture model.



Note: 1 “Immigration should be facilitated” (pro) and 11 “Immigration should be restricted” (contra)

Fig. 2 Distribution of party perceptions on the immigration issue

4.3 Predictors for Issue Placements and Issue Uncertainty

As outlined above, the approach can model the placement and uncertainty structure of political perceptions by covariates x_i and w_i . For both components, we use the same set of predictors we describe next.

4.3.1 Political Sophistication

The first set of explanatory variables accounts for cognitive processes or information costs and relates to the concept of political sophistication (e.g., Luskin 1987, 1990; Delli Carpini and Keeter 1993). Going back to Downs (1957), it is frequently argued that citizens who possess lower information costs tend to be more informed about the positions parties offer on central policies. Therefore, individuals with higher levels of political information are presumed to be less uncertain about party platforms. To operationalize the concept of political sophistication and to identify segments that might exhibit different response styles due to uncertainty or place parties into a particular direction, we explore a subjective and an objective measure: the strength of political interest, and political knowledge.

The level of political interest is usually measured by relying on respondents' self-reports. The 2017 German election study includes a question in which respondents were asked to state their level of political interest on a five-point scale. We recoded the variable so that one gives the response "not interested at all" and five "very interested". To measure the respondents' level of political knowledge we rely on factual knowledge questions with right or wrong answers. Several studies have shown that factual political knowledge questions present good empirical indicators for the concept of political knowledge (Luskin 1987, 1990; Delli Carpini and Keeter 1993).³ Based on the replies to five questions, we generated an additive knowledge score in which for each correct answer a value of one is assigned, whereas wrong answers and "don't know/no answer" responses give a value of zero. The first two questions concern the German electoral system.⁴ In addition, the respondents were confronted with pictures showing three politicians, and they were asked to state the party each politician belongs to.⁵ The answers are

³For a recent comparative assessment, see Rapeli (2013).

⁴"Which one of the two votes is decisive for the relative strengths of the parties in the German parliament?"; "What is the percentage of the second vote a party needs to be able to definitely send delegates to the German parliament?".

⁵These politicians are Martin Schulz (SPD), Katrin Göring-Eckardt (Greens), and Christian Lindner (FDP).

aggregated by counting the number of times a respondent correctly answered all five questions, resulting in a six-categorical variable running from zero to five (0 none correct, 5 all answers correct).

4.3.2 Party Attributes

One might also expect that respondents are more confident about where to place the party due to the relationship they have established with the respective party. The second set of covariates intends to capture the relationship between the respondent and the party to be located. One might expect that a long-standing leaning toward a party influences both where to place the party and the response patterns. So that those respondents who identify themselves with the party to be located are more certain about the position the party offers. On the contrary, when respondents do not identify themselves with the respective party, specific response styles due to uncertainty might be likely to occur. The same argument might apply to the sympathy of the parties' candidates. Party identification is a dummy variable with one indicating that the respondent identifies with the party to be placed and zero otherwise (i.e., no party identification or identification with any other party). For each of the seven parties, we generated such a dummy variable. As candidate evaluations, we consider feeling thermometers on eleven-point scales (1 very negative; 11 very positive).⁶

4.3.3 Issue Importance

Also, the issue itself might influence the respondents' perceived party placements and uncertainty. Respondents who consider the policy as important might have a clearer understanding of where to place the parties. When the issue is of personal importance, the respondent might have considered in more detail what the parties actually offer on it. Therefore, we might expect that as the level of personal issue importance increases, the respondents show less uncertainty response patterns. To identify the level of importance respondents ascribe to the immigration issue, we employ a common measure, self-reports. The 2017 German election study includes a question in which respondents are asked to state the importance of the immigration issue on a five-point scale running from "not at all important" to "very important".⁷

⁶The candidates are Angela Merkel for the CDU, Horst Seehofer for the CSU, Martin Schulz for the SPD, Christian Lindner for the FDP, Cem Özdemir for the Greens, Sahra Wagenknecht for the Left, and Frauke Petry for the AfD.

⁷Corresponding question: "How important is this issue to you personally?".

4.3.4 Standard Demographics

Finally, the models account for the effects of standard demographics, including age and gender. Since East/West Germany constituted for a long time a major explanation for differences in public opinion in Germany, the models also control for this east-west divide. The variables are coded as follows: Age: centered around the sample mean, measured in decades; Gender: 1 (female), 0 (male); Former West/East Germany: 1 (West Germany), 0 (East Germany). As with the party placements, we excluded missing values on all these variables.

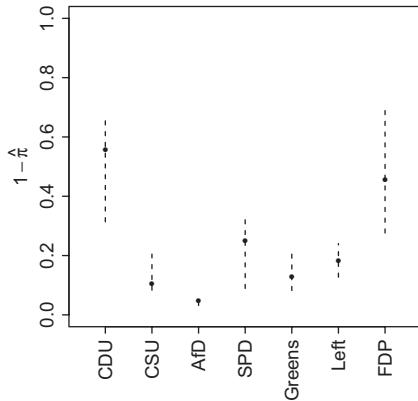
5 Empirical Results

For each party, we specified a separate BetaBin model. The models are estimated with the EM-Algorithm as described in Tutz and Schneider (2018). The result presentation is divided into three parts: We begin by examining the role uncertainty plays in the party perceptions on the immigration issue. Then, we present the estimates for the placement and the uncertainty part of the model. Finally, we systematically compare the proposed BetaBin model with the traditional cumulative model without uncertainty component based on performance measures.

5.1 The Role of Uncertainty in Perceived Party Platforms

Let us first examine the mixture probability that measures the importance of the uncertainty part in the mixture model. $1 - \hat{\pi}$ is an estimate for the weight of the uncertainty component and can be interpreted as a measure of how clear or unambiguous the respondents perceive the party positions. The weight takes values between zero and one. A value of zero represents the traditional cumulative model without mixture and indicates the absence of any response styles, i.e., no tendency to the middle of the scale, no tendency to extremes, and no random party perception. A value of one gives a model without placement structure, i.e., no structure in placing the parties on the policy scales is detected. The higher the value of $1 - \hat{\pi}$, the stronger the uncertainty.

Figure 3 displays the estimated $1 - \hat{\pi}$ values for the seven German parties. The dotted lines correspond to the 2.5% and 97.5% bootstrap quantiles. As expected, we detect the weakest uncertainty weight (0.05) for the location of the AfD. This indicates that respondents expressed a clear preference where to place the AfD on the immigration issue. Regarding the two Christian Democratic parties, the



Note: Dotted lines correspond to the 2.5% and 97.5% bootstrap quantiles.

Fig. 3 Importance of the uncertainty component ($1 - \hat{\pi}$)

respondents exhibit much more uncertainty, 0.56 and 0.11, respectively. A higher level of uncertainty for these two right-of-center parties seems plausible because of the internal divisions on the immigration issue. The chancellor and leader of the CDU Angela Merkel, Thomas de Maizi  re (Minister of the Interior and member of the CDU), and Horst Seehofer (leader of the CSU) expressed quite different opinions about migrants and refugees. Thus, the respondents are more uncertain about both positions. However, the CSU could offer a much clearer position, as suggested by the considerable difference in uncertainty between the CDU and CSU. The behavior on the side of the parties is reflected in the uncertainty weights we estimate. Our model also estimates a large uncertainty weight for the FDP. Apparently, respondents show enormous difficulties in placing both the CDU and the FDP on the immigration issue, whereas the strength of uncertainty is in the middle range for the remaining parties, except for the AfD. Note that the uncertainty weight illustrates the strength of the uncertainty component regardless of whether the perceived party position is pro or contra immigration.

5.2 Party Placements and Response Style Effects

Tables 1 and 2 give the results for the placement and uncertainty components of the models. The estimates for the placement part ($\hat{\gamma}$) are displayed at the top, the estimates for the uncertainty response style effects ($\hat{\alpha}$) at the bottom. We report

Table 1 Parameter estimates of BetaBin model I

	CDU			CSU			AfD		
	Estimate	BS.2.5	BS.97.5	Estimate	BS.2.5	BS.97.5	Estimate	BS.2.5	BS.97.5
<i>Placement part</i>									
Political interest	-0.281	-0.454	0.049	-0.131	-0.278	0.018	-0.147	-0.299	0.006
Political knowledge	-0.234	-0.374	-0.064	*	-0.340	-0.462	-0.233	*	-0.079
Party identification	0.423	-0.120	0.909	-0.771	-1.712	-0.127	*	0.055	-0.560
Candidate images	-0.183	-0.372	-0.065	*	-0.164	0.113	0.238	*	0.250
Issue importance	-0.091	-0.521	0.212	-0.068	-0.206	0.067	-0.127	-0.277	0.016
Age	-0.167	-0.285	-0.053	*	-0.194	-0.283	-0.150	*	-0.145
Gender	-0.414	-0.748	0.006	-0.218	-0.468	0.024	-0.201	-0.471	0.050
West Germany	-0.044	-0.463	0.314	-0.050	-0.245	0.204	0.183	-0.068	0.468

(Fortsetzung)

Table 1 (Fortsetzung)

	CDU			CSU			AfD		
	Estimate	BS.2.5	BS.97.5	Estimate	BS.2.5	BS.97.5	Estimate	BS.2.5	BS.97.5
<i>Uncertainty part</i>									
Political interest	0.127	-0.293	0.611	-0.123	-3.236	1.381	0.843	-2.930	6.211
Political knowledge	0.385	0.191	1.183	*	1.038	0.043	3.974	*	-0.417
Party Identification	0.230	-0.381	3.891	10.80	-1.063	255.78	1.176	-10.38	8.248
Candidate images	0.163	0.071	0.448	*	0.254	-0.697	1.761	0.545	0.768
Issue importance	-0.264	-1.574	0.029	-0.485	-4.050	1.577	-1.088	-5.934	1.100
Age	-0.192	-0.572	-0.069	*	-0.378	-1.872	0.633	0.202	-1.091
Gender	-0.049	-0.845	0.704	0.818	-2.451	5.068	-0.365	-6.158	4.866
West Germany	0.323	-0.532	1.083	-0.009	-3.813	3.082	-1.763	-8.680	-0.979
$1 - \hat{\pi}$	0.558	0.314	0.657	*	0.106	-0.083	0.205	*	0.047
N		1949				1872		1715	

Note: Cut points of the placement part and intercept of the uncertainty part are not displayed

* An estimate is considered as significant at the 5%-level when the bootstrap confidence intervals cover the estimate but not zero

Table 2 Parameter estimates of BetaBin model II

	SPD			Greens			Left			FDP		
	Estimate	BS.2.5	BS.97.5									
<i>Placement part</i>												
Political interest	0.347	0.185	0.544 *	0.426	0.310	0.588 *	0.405	0.276	0.584 *	0.039	-0.353	0.309
Political knowledge	0.100	-0.004	0.204	0.266	0.180	0.380 *	0.240	0.140	0.333 *	-0.214	-0.545	0.047
Party identification	0.211	-0.039	0.511	0.375	0.034	0.681 *	-0.195	-0.525	0.233	0.396	-0.860	1.505
Candidate Images	-0.080	-0.156	-0.022 *	-0.184	-0.254	-0.118 *	-0.010	-0.068	0.042	0.135	-0.004	0.613
Issue importance	0.014	-0.133	0.246	0.095	-0.061	0.281	0.085	-0.049	0.240	-0.325	-1.020	-0.047 *
Age	-0.135	-0.198	0.054	-0.015	-0.074	0.050	-0.187	-0.253	-0.127 *	-0.090	-0.279	0.023
Gender	-0.264	-0.493	0.001	-0.134	-0.368	0.097	-0.442	-0.677	-0.231 *	-0.067	-0.616	0.667
West Germany	-0.257	-0.572	-0.014 *	-0.342	-0.603	-0.095 *	-0.870	-1.144	-0.615 *	0.249	-0.416	0.857

(Fortsetzung)

Table 2 (Fortsetzung)

	SPD			Greens			Left			FDP		
	Esti-	BS.2.5	BS.97.5	Esti-	BS.2.5	BS.97.5	Estimate	BS.2.5	BS.97.5	Esti-	BS.2.5	BS.97.5
<i>Uncertainty part</i>												
Political interest	0.116	-0.938	2.048		-0.259	-2.110	3.162	0.638	-0.541	2.482	-0.468	-1.908
Political knowledge	0.410	0.130	2.724	*	0.840	-0.683	2.598	1.325	0.321	2.129	*	0.530
Party identification	-0.338	-3.615	2.891		-3.867	-7.972	13.73	-1.065	-6.663	3.798	7.921	-1.045
Candidate images	0.150	-1.061	1.105		0.440	-0.767	1.251	0.895	0.343	1.564	*	0.402
Issue importance	-1.156	-4.253	1.466		-3.505	-4.907	1.490	-3.818	-5.044	-0.870	*	-0.022
Age	-0.461	-1.740	0.005		-0.728	-1.853	0.239	-0.234	-1.219	0.480	-0.234	-0.710
Gender	-0.546	-3.498	3.071		5.846	0.505	8.671	*	0.341	-2.642	2.402	-0.102
West Germany	0.778	-2.612	3.635		1.927	-1.805	8.687	-4.534	-7.096	-1.843	*	0.736
$1 - \hat{\pi}$	0.249	0.089	0.336	*	0.128	0.082	0.205	*	0.181	0.127	0.240	*
N		1813					1621			1527		1387

Note: Cut points of the placement part and intercept of the uncertainty part are not displayed

for each effect the 2.5% and 97.5% quantiles of 500 non-parametric bootstrap samples. An estimate is considered as significant at the 5%-level when the bootstrap confidence intervals cover the estimate but not zero. The estimates for the preference part ($\hat{\gamma}$) are interpreted in the following way. Positive estimates indicate that lower categories become more likely meaning that the respondents tend to place the party toward pro immigration. Negative values indicate a tendency to locate the party in a higher category on the issue scale corresponding with contra-immigration stances.

We observe that the effects of political sophistication vary between the different parties. When the respondents' political interest increases, they tend to locate the SPD, the Greens and the Left closer toward pro-immigration stances. Higher political knowledge leads to a location of both Christian Democrat parties closer to contra-immigration positions, whereas the opposite is the case for the Greens and the Left. These results suggest that for the majority of the parties political sophistication does provide an explanation of placement structures. Regarding our second set of predictors, we find that candidate images are significant for all parties, except for the Left and FDP. However, the estimates differ in direction. Whereas an increase in candidate sympathy leads respondents to locate the CSU and the AfD toward pro-immigration positions, it yields placements closer to contra-immigration stances for the CDU, SPD, and the Greens. By contrast, party identification only impacts on the placement of two parties, namely the CSU and the Greens. When respondents identify themselves with the Greens, they show a tendency to locate the party closer at pro-immigration stances. In the case of the CSU, the effect is negative so that CSU-party identifiers tend to place their party closer to contra-immigration positions. Concerning issue importance, there are almost no significant findings. By contrast, the demographics exhibit some interesting findings. We observe significant negative age effects for several parties. This implies that the older a person, the stronger the tendency to ascribe the parties more contra-immigration positions. In addition, females show a tendency to locate the Left closer at contra-immigration stances, whereas respondents based in former West Germany tend to shift the SPD, the Greens and the Left toward contra-immigration positions.

The uncertainty component contains the estimates for the shape of the uncertainty distribution (determined by $\hat{\alpha}$), displayed at the bottom of Tables 1 and 2. Here, positive values indicate a tendency to the middle categories, whereas negative values suggest a tendency to the extremes of the scales. As in the placement structure, measures of political sophistication also show some significant effects on the response style dimension. The higher the political knowledge, the more respondents tend to locate the CDU, the CSU, the SPD, and the Left in middle

categories. By contrast, political interest does not seem to influence the response styles. Examining the predictors that relate to the relationship between the respondent and the party to be located, we find the strongest effects for candidate images, whereas party identification shows no significant effects. We observe that an increased satisfaction with the candidate leads to a tendency to locate the CDU, the AfD, and the Left in middle categories. Also, issue importance only marginally impacts on response styles. Surprisingly, most demographic variables do not exhibit any significant effects. We observe that older people tend to locate the CDU toward more extreme positions, females show a tendency to place the Greens in middle categories, and respondents from former West Germany ascribe extreme positions to both the AfD and the Left.

5.3 Model Comparisons

Finally, we demonstrate that the mixture model outperforms the traditional cumulative model based on performance measures. In Table 3, we compare the performances of the proposed BetaBin model with the cumulative model without uncertainty component. We measure model performance by the Log-Likelihood and the AIC. The latter is defined by

$$AIC = -2l(\hat{\theta}) + 2m,$$

where m is the number of model parameters and $l(\hat{\theta})$ is the log-likelihood function computed at the maximum of the estimated parameter vector $\hat{\theta}$. We see that

Table 3 Model comparisons based on performance measures

Model	N	LogL		AIC	
		Mixture	Cumulative	Mixture	Cumulative
CDU	1949	-4314.096	-4370.631	8684.191	8777.262
CSU	1872	-3926.739	-3941.070	7909.478	7918.140
AfD	1715	-1706.051	-1723.742	3468.102	3483.484
SPD	1813	-3620.930	-3637.966	7297.860	7311.933
Greens	1621	-3153.747	-3171.438	6363.494	6378.875
Left	1527	-3299.113	-3326.538	6654.227	6689.075
FDP	1387	-2652.197	-2675.003	5360.394	5386.006

the mixture model improves all considered performance measures as compared to the cumulative model without any uncertainty component. All AIC values are lower for the mixture than for the pure cumulative model. While the pure cumulative model is based on 18 parameters (10 intercepts and 8 covariables), the mixture model is based on a total of 28 parameters: 18 parameters for the placement part, which is identical with the pure cumulative model, 9 parameters for modeling the shape of the uncertainty distribution (1 intercept and 8 covariables) and 1 parameter to estimate the mixture weight π . Even though the mixture model is much more complex, the performance measures indicate that it yields not only a better likelihood but also to a better model fit measured by AIC. Since the number of observations differs among parties, the values can only be compared across the different models but not across parties.

6 Discussion and Outlook

Political perceptions play an important role in the decision-making process. In this paper, we have applied a special mixture model, the so-called BetaBin model, to the perception of party placements on ordinal policy scales. The model consists of two components, a placement part and an uncertainty part. The latter enables us to model response styles to the middle categories as well as to extreme categories. For the placement part, a cumulative model is used, and for the uncertainty part, we rely on a restricted beta-binomial distribution. We applied the model to the immigration issue in the occasion of the 2017 German national election. Our results demonstrate that the respondents' perceptions of most parties on the immigration issue are strongly influenced by uncertainty. We detect the lowest uncertainty in locating the AfD and the highest in placing the FDP and the CDU. Regarding the predictors we examined, we find that particularly political sophistication and candidate images influence both where to place the party and the uncertainty response patterns. Especially interesting are also the age effects we detect. The older the respondents, the more they tend to locate the parties toward the contra-immigration pole. Finally, our model outperforms traditional cumulative models without uncertainty structure based on model performance.

Next steps will be to apply the model to other policy scales and contexts. Another interesting aspect would be to examine all parties simultaneously in a multivariate model. In the proposed model, we use covariates to model the uncertainty structure and the placement part, but not to model the mixture weights π . There are other approaches (e.g., Tutz et al. 2017) which use covariates in the placement part and the mixture weights, but not in the uncertainty part. Including

covariates in all three components of the mixture model may lead to identifiability issues which have not been discussed yet. In future research, we also intend to develop a voter choice model that relies on the model of survey responses we proposed here. A voter choice model that is based on the party placement and uncertainty estimates of the mixture model will then add to our understanding of how uncertainty impacts on policy-oriented decision making and its electoral consequences. Then it will also be possible to examine and test for behavioral implications of uncertainty in political perceptions.

Acknowledgements We would like to thank the two anonymous reviewers and the editors for their very valuable comments and suggestions. We also thank Hannah Reitz for assistance in preparing the tables.

References

- Abou-Chadi, Tarik. 2015. Niche party success and mainstream party policy shifts – how green and radical right parties differ in their impact. *British Journal of Political Science* 46 (2): 417–436.
- Aldrich, John H., Richard G. Niemi, George Rabinowitz, and David W. Rohde. 1982. The measurement of public opinion about public policy: A report on some new issue question formats. *American Journal of Political Science* 26 (2): 391–414.
- Alonso, Sonia and Sara Claro Da Fonseca. 2012. Immigration, left and right. *Party Politics* 18 (6): 865–884.
- Alvarez, R. Michael. 1999. *Information and Elections*. Ann Arbor: University of Michigan Press.
- Alvarez, R. Michael and John Brehm. 1995. American ambivalence towards abortion policy: Development of a heteroskedastic probit model of competing values. *American Journal of Political Science* 39 (4): 1055–1082.
- Alvarez, R. Michael and John Brehm. 1997. Are Americans ambivalent towards racial policies? *American Journal of Political Science* 41 (2): 345–374.
- Alvarez, R. Michael and John Brehm. 1998. Speaking in two voices: American equivocation about the internal revenue service. *American Journal of Political Science* 42 (2): 418–452.
- Alvarez, R. Michael and John Brehm. 2002. *Hard Choices, Easy Answers: Values, Information, and American Public Opinion*. Princeton: Princeton University Press.
- Alvarez, R. Michael and Charles H. Franklin. 1994. Uncertainty and political perceptions. *The Journal of Politics* 56 (3): 671–688.
- Art, David. 2011. *Inside the Radical Right: The Development of Anti-Immigrant Parties in Western Europe*. New York: Cambridge University Press.

- Bale, Tim. 2003. Cinderella and her ugly sisters: The mainstream and extreme right in Europe's bipolarising party systems. *West European Politics* 26 (3): 67–90.
- Bale, Tim. 2008. Turning round the telescope. Centre-right parties and immigration and integration policy in Europe. *Journal of European Public Policy* 15 (3): 315–330.
- Bale, Tim, Christoffer Green-Pedersen, Andre Krouwel, Kurt Richard Luther, and Nick Sitter. 2010. If you can't beat them, join them? Explaining social democratic responses to the challenge from the populist radical right in Western Europe. *Political Studies* 58 (3): 410–426.
- Bartels, Larry M. 1986. Issue voting under uncertainty: An empirical test. *American Journal of Political Science* 30 (4): 709–728.
- Baumgartner, Hans and Jan-Benedict Steenkamp. 2001. Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research* 38 (2), 143–156.
- Berinsky, Adam J. and Jeffrey B. Lewis. 2007. An estimate of risk aversion in the U.S. electorate. *Quarterly Journal of Political Science* 2 (2): 139–154.
- Campbell, Angus, Philip E. Converse, E. Warren Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago: University of Chicago Press.
- 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 (2): 426–448.
- De Vries, Catherine E. and Marco R. Steenbergen. 2013. Variable opinions: The predictability of support for unification in European mass publics. *Journal of Political Marketing* 12 (1): 121–141.
- D'Elia, Angela and Domenico Piccolo. 2005. A mixture model for preference data analysis. *Computational Statistics & Data Analysis* 49 (3): 917–934.
- Delli Carpini, Michael X. and Scott Keeter. 1993. Measuring political knowledge: Putting first things first. *American Journal of Political Science* 37 (4): 1179–1206.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Enelow, James M. and Melvin J. Hinich. 1981. A new approach to voter uncertainty in the Downsian spatial model. *American Journal of Political Science* 25 (3): 483–493.
- Gill, Jeff. 2005. An entropy measure of uncertainty in vote choice. *Electoral Studies* 24 (3): 371–392.
- Givens, Terri E. 2005. *Voting Radical Right in Western Europe*. New York: Cambridge University Press.
- Harbers, Imke, Catherine E. De Vries, and Marco R. Steenbergen. 2013. Attitude variability among Latin American publics: How party system structuration affects left/right ideology. *Comparative Political Studies* 46 (8): 947–967.
- Harvey, A. C. 1976. Estimating regression models with multiplicative heteroscedasticity. *Econometrica* 44 (3): 461–465.
- Iannario, Maria and Domenico Piccolo. 2016a. A comprehensive framework of regression models for ordinal data. *Metron* 74 (2): 233–252.
- Iannario, Maria and Domenico Piccolo. 2016b. A generalized framework for modelling ordinal data. *Statistical Methods & Applications* 25 (2): 163–189.
- Ignazi, Piero. 2003. *Extreme Right Parties in Western Europe*. Oxford: Oxford University Press.

- Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton: Princeton University Press.
- Ivarsflaten, Elisabeth. 2008. What unites right-wing populists in Western Europe? Reexamining grievance mobilization models in seven successful cases. *Comparative Political Studies* 41 (1): 3–23.
- Kedar, Orit. 2005a. How diffusion of power in parliaments affects voter choice. *Political Analysis* 13 (4): 410–429.
- Kedar, Orit. 2005b. When moderate voters prefer extreme parties: Policy balancing in parliamentary elections. *The American Political Science Review* 99 (2): 185–199.
- Kitschelt, Herbert and Anthony McGann. 1995. *The Radical Right in Western Europe: A Comparative Analysis*. Ann Arbor: University of Michigan Press.
- Luskin, Robert C. 1987. Measuring political sophistication. *American Journal of Political Science* 31 (4): 856–899.
- Luskin, Robert C. 1990. Explaining political sophistication. *Political Behavior* 12 (4): 331–361.
- Mauerer, Ingrid. 2016. A party-varying model of issue voting. A cross-national study (doctoral dissertation). *University of Munich (LMU), Germany*.
- Mauerer, Ingrid, Paul W. Thurner, and Marc Debus. 2015. Under which conditions do parties attract voters' reactions to issues? Party-varying issue voting in German elections 1987–2009. *West European Politics* 38 (6): 1251–1273.
- McLachlan, Geoffrey J. and David Peel. 2000. *Finite Mixture Models*. New York: Wiley.
- Meguid, Bonnie M. 2005. Competition between unequals: The role of mainstream party strategy in niche party success. *The American Political Science Review* 99 (3): 347–359.
- Mudde, Cas. 2007. *Populist Radical Right Parties in Europe*. Cambridge: Cambridge University Press.
- Norris, Pippa. 2005. *Radical Right: Voters and Parties in the Electoral Market*. New York: Cambridge University Press.
- Pardos-Prado, Sergi, Bram Lancee, and Iñaki Sagarzazu. 2014. Immigration and electoral change in mainstream political space. *Political Behavior* 36 (4): 847–875.
- Rapeli, Lauri. 2013. *The Conception of Citizen Knowledge in Democratic Theory*. Basingstoke: Palgrave Macmillan.
- Rossteutscher, Sigrid, Harald Schoen, Rüdiger Schmitt-Beck, Bernhard Wessels, Christoph Wolf, Ina Bieber, Lars-Christopher Stövsand, and Melanie Dietz. 2017. Pre-election cross section (GLES 2017). ZA6800 Data file Version 3.0 Cologne: GESIS Data Archive.
- Rozenas, Arturas. 2013. Inferring ideological ambiguity from survey data. In *Advances in Political Economy: Institutions, Modelling and Empirical Analysis*, eds. Norman Schofield, Gonzalo Caballero, and Daniel Kselman, 369–382. Berlin, Heidelberg: Springer.
- Shepsle, Kenneth A. 1972. The strategy of ambiguity: Uncertainty and electoral competition. *The American Political Science Review* 66 (2): 555–568.
- Stoetzer, Lukas F. 2017. A matter of representation: Spatial voting and inconsistent policy preferences. *British Journal of Political Science* Advance online publication.
- Tomz, Michael and Robert P. Van Houweling. 2009. The electoral implications of candidate ambiguity. *American Political Science Review* 103 (1): 83–98.
- Tutz, Gerhard. 2012. *Regression for Categorical Data*. Cambridge: Cambridge University Press.

- Tutz, Gerhard and Micha Schneider. 2017. Mixture models for ordinal responses with a flexible uncertainty component. Technical Report 203, Department of Statistics, University of Munich (LMU), Germany.
- Tutz, Gerhard and Micha Schneider. 2018. Flexible uncertainty in mixture models for ordinal responses. *Journal of Applied Statistics*, <https://doi.org/10.1080/02664763.2018.155574>.
- Tutz, Gerhard, Micha Schneider, Maria Iannario, and Domenico Piccolo. 2017. Mixture models for ordinal responses to account for uncertainty of choice. *Advances in Data Analysis and Classification* 11 (2): 281–305.
- Van Vaerenbergh, Yves and Troy D. Thomas. 2013. Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research* 25 (2): 195–217.



Risk Attitudes, Gender, and Risk Behavior: Evidence from Two Laboratory Experiments

Christine Prokop

Abstract

This study analyzes whether self-reported attitudes in economic risk taking predict experimentally revealed risk behavior, and how gender moderates the relationship between both measures. Prior research often finds women reporting higher risk averse attitudes than men and showing more risk averse behavior in observational or experimental studies. This study analyzes observations from 369 students who participated in two laboratory experiments and answered a survey about their risk preferences. The findings show that risk attitudes are not likely to predict risk behavior directly, but being female predicts risk averse behavior robustly. Most interestingly, the analyses show that in the experiments, women behave consistently to their self-reported risk attitudes, but men do not. Methodological and practical implications are briefly discussed.

I am very grateful for the support and valuable comments on this paper from Markus Tepe and Michael Jankowski. I wish to thank the participants of the 2018 Annual Meeting of the DVPW working group on decision theory for the fruitful discussion on a prior version of this paper, as well as the two anonymous reviewers for their thorough comments, which further helped to improve this work.

C. Prokop (✉)

Fakultät für Bildungs- und Sozialwissenschaften, Oldenburg, Germany
E-Mail: christine.prokop@uni-oldenburg.de

1 Introduction

Survey research often suggests that women perceive themselves as more risk averse than men. Experimental research reveals that women decide more risk aversely than men. It is unclear, however, whether the observed effects in risk behavior correlate with the levels of self-reported risk attitudes or whether there is still a gender effect in risk behavior once the self-reported risk attitudes are taken into account.

Decision making under risk has a long tradition in economic research. Profound knowledge about risk attitudes and their behavioral impact, however, is important in the political and social arenas as well. Many political and public administrative decisions touch risky financial topics. The practical implications of gender differences, for instance in leadership behavior (e.g. O'Brien 2015; Meier et al. 2006), are a growing field in political science and public administration research. In contrast to the popular understanding of risk as a chance of some negative event, this study utilizes the classic risk definition of Knight (1921). Knight defines risk as measurable uncertainty that arises when the actual outcome of an event can deviate from the expected result. Laboratory experiments with monetary incentives allow to measure this type of risk behavior.

Generally, a congruence between surveyed and experimentally measured risk preferences can be doubted (Ajzen and Fishbein 1977). While survey studies elicit self-perceived risk *attitudes*, risk preferences observed using monetarily incentivized experiments measure risk *behavior*. Surveyed attitudes can be biased by self-stereotyping (Daruvala 2007) or strategic motives (e.g., Holt 2007; Gintis 2000). Laboratory experiments provide the highest internal validity toward the causal mechanism affecting risk behavior due to the highly controlled environmental conditions.

This study uses data from two well-established laboratory experimental vehicles—the first-price sealed-bid auction (Auction) and the compound lottery choice treatment (Lottery)—and estimates the effect of self-reported risk attitudes on the revealed risk behavior. The purpose of this study is to elicit whether gender differences can be found in behavioral consequences of self-reported risk attitudes. Note that this study addresses a correlational rather than a causal relationship between individual attitudes and behavior. The research question addressed in this work is: Are women more likely to behave consistently to their stated risk attitudes than men?

In the following, a brief overview over the relevant risk literature is given. After introducing the method and experimental procedures, the empirical analysis and its results are presented. The last chapter summarizes the study and discusses its implications.

2 Literature Review on Gender Risk Preferences and Hypothesis

2.1 Risk Attitudes and Risk Behavior

Risk preferences refer either to risk attitudes or to risk behavior. From a theoretical perspective, the relationship between the two aspects depends on the attitudinal and behavioral entities (Ajzen and Fishbein 1977). The degree to which these entities match predicts the strength of the correlation between attitudes and behavior. From a methodological perspective, risk attitudes are stated, and risk behavior are revealed risk preferences. Different methods apply to elicit information about one or the other risk preference. Attitudes are typically measured using surveys; data on behavior can either be observed in actual life situations or measured with experimental instruments. Now it is compelling to assume that stated risk attitudes can fail to fully match risk behavior in situations with real (monetary) consequences, like experimental tasks, because they refer to hypothetical rather than factual risk situations.

This sketches the frame why surveyed risk preferences sometimes predict revealed risk behavior and sometimes fail to do so. One important methodological asset is the use of monetary incentives to elicit measures on risk behavior. Quasi-experimental designs that, for instance, forego monetary incentives or expose their participants to hypothetical scenarios, are more likely to elicit risk attitudes than risk behavior due to ‘cheap talk’ or social desirability bias (Maccoby and Maccoby 1954). This study uses two monetary incentivized laboratory experiments to measure risk behavior and a post-experiment survey on economic risk attitudes. The risk survey measures are supposed to predict risk behavior given prior empirical findings reported by Dohmen et al. (2011), Kraus (1995) and Lönnqvist et al. (2015).

It is important to note that in psychological as well as in economic literature, the stability of individual risk preferences is questioned. While (risk) attitudes can vary over time or context (e.g. Bohner and Dickel 2011; Dohmen et al. 2011; Lönnqvist et al. 2015), risk behavior is likely to respond to experimental treatments (e.g. Berg et al. 2005; Levin et al. 2007), the size of incentives (e.g. Holt and Laury 2002), or to differences in framing (e.g. Kahneman 1992; Levin et al. 1988). Some studies, in contrast, indicate that risk behavior reveals intrapersonal stability over time and elicitation method (e.g. Harrison et al. 2005; Pedroni et al. 2017). However, the potential intrapersonal variance of risk preferences is not confounding this analysis, because of its strong internally valid research design. This study employs a

strictly controlled laboratory experimental setting with a neutral framing and monetary incentives to minimize and control for external effects on the risk measures. Further, experimental task and survey measures are separated by unrelated output stages and survey questions to avoid spillover effects. Full random assignment to the treatments further allows for between-subject comparison on a high degree of internal validity. Considering the theoretical reasoning, the first hypothesis tested in this paper reads as follows:

H1 Self-reported risk attitudes predict experimental risk behavior.

2.2 Prior Research on Gender Risk Preferences

This study analyzes the moderating impact of gender on the attitudes-behavior relationship. Thus, as a starting point, a brief overview of prior findings in gender risk preferences is provided in the following. Byrnes et al. (1999) provide an overview over psychological research about gender risk attitudes. Their meta-study of 150 studies on gender differences in risk taking includes studies using self-reported or observational data, and studies employing hypothetical gambles. The overall findings are that women are in most situations more risk averse than men. The magnitude of this gender difference varies depending on the context. Economic research provides similar findings using survey data (e.g. Halek and Eisenhauer 2001; Morin and Suarez 1983). Dohmen et al. (2011) analyze self-reported risk attitudes using a representative large scale sample in Germany. These risk attitudes include a measure for the willingness to take risks in general and in five sub dimensions (car driving, financial matters, sports and leisure, career). Their survey findings indicate that among other correlations, being female is positively correlated with higher risk aversion.

A broad range of economic literature provides support for women deciding more risk averse in experimental environments. Several studies find women deciding more risk aversely in investment decision tasks (e.g. Charness and Gneezy 2012; Cohn et al. 2017; Wärneryd 1996). Choosing from a set of gambles typically elicits more risk averse behavior among women as well (e.g. Eckel and Grossman 2002; Eckel and Grossmann 2008; Schubert et al. 1999). The lottery choice task employed by Fehr-Duda et al. (2006) elicits that men and women prefer the same value schemes but show a different sensibility to probabilities, which is why women appear more risk averse. Fellner and Maciejovsky (2007) find that

women play more risk averse in lotteries and make fewer offers and less trades in a market scenario. The studies of Holt and Laury (2002), Harrison and Rutström (2008), and Bruhin et al. (2010) provide support to the female risk aversion using lotteries as well. The review studies by Croson and Gneezy (2009) and Charness et al. (2013) compare studies using a range of experimental methods, like the balloon analogue risk task, decision tasks, multiple price list method, or gambles, and typically find women acting more risk averse than men or find no difference. The natural experiment by Kelly and Lemke (2015) using data from the TV show Cash Cab shows that women play less risky and consider available information more often than men. Powell and Ansic (1997) show that the observed female risk aversion is not affected by the experimental environment, like by framing, or task familiarity. None of these studies indicates that men behave more risk averse than women. Based on these empirical findings, the second hypothesis is:

H2 Women show more risk averse behavior in the experimental settings.

2.3 Gender Effect on Risk Attitudes-Behavior Relationship

As discussed above, according to the Ajzen and Fishbein (1977) framework risk attitudes translate stronger into risk behavior the more congruent the attitudinal and behavioral entities are. In other words, how well self-perceived risk attitudes predict risk behavior depends on the ability to assess the own attitudes sincerely. Revealing more about the specific roots of female risk aversion, however, goes beyond the scope of this study, but it is argued that gender specific differences in the assessment and evaluation of risk lead to gender differences in risk preferences. Nevertheless, the empirically observable gender effects in risk preferences indicate a fundamental difference between men and women. A possible explanation could be a gender specific subjective perception of risk (Finucane et al. 2000) or risk related gender specific role-models (Lemaster and Strough 2014). Moreover, emotions, the level of overconfidence or framing are discussed in literature (Croson and Gneezy 2009). Socialization effects and self-stereotyping might lead to higher reported risk levels through social desirability effects among individuals who evaluate risk taking as positive characteristic and vice versa. How risk behavior is affected by these self-perceptions separated by gender is object to the last hypothesis:

H3 Gender moderates the predictive power of risk attitudes on risk behavior.

3 Methods and Data

3.1 Risk Attitudes

A survey to elicit sociodemographic characteristics and risk attitudes was provided to the participants after the experiment was finished. Conducting the experimental task first is necessary to avoid framing the participants by asking questions toward their risk attitudes prior to a risky decision task. Between the last entry stage and the risk survey, the participants were given information like the experimental outcome and sociodemographic and not risk taking-related survey questions. Participants thus responded to the risk questionnaire a reasonable amount of time after the game situation, so that the game should not affect the self-assessment.

The risk dimensions used in this survey stem from the German Socio-Economic Panel (GSOEP 2004). The exact wording of the question (translated into English) is given in Appendix Table 5. Participants were asked about their willingness to take risk in the domains of driving, financial matters, sports and leisure, career, health, and trust in other people (see Dohmen et al. 2011). On a scale ranging from 0 to 10, higher values imply more risk averse attitudes. To keep the analysis as simple as possible, only the risk domains related to economic risk taking are used for analysis, i.e. risk attitudes toward financial matters and career, which are combined to a single risk attitudes variable. All risk domains are included in robustness analyses to allow for comparing and contrasting the findings.

3.2 Auction

The first experimental instrument to measure risk behavior in this study is a first-price sealed-bid auction (Berg et al. 2005; Milgrom and Weber 1982). During the game, participants gain points that are converted into Euro and paid in cash at the end of the experiment. The auction is played with two bidders—one participant and one computerized bidder.

The instructions cover that the computer's bid is an equiprobable number between one and 50 and the participant's endowment in points is in each round a randomly chosen number between one and 100.¹ To win an auction, the participant's

¹The computerized bidder plays a risk neutral bidding strategy, which is not revealed to the participants.

bid has to be higher than the bid of the computer. Ties benefit the computerized bidder (for the entire instructions see Appendix). In case of winning an auction, the winning points are calculated as the residual amount of the initial endowment minus the amount of bedded points in each round. The winning points per period are added to the player's account.² The auction was repeated ten times in total.

To avoid testing the participant's mathematical skills, a calculator is provided. Before confirming the bid, participants are allowed to calculate the expected winning points, the winning probability, and the expected profit (expected profit=expected winning points * winning probability) as often as they wish. After all players have confirmed their bids, an output screen displays the valuation, the confirmed bid, computer's bid, the winning bid, the winning points, and in addition a note whether the player did or did not win the auction as well as each of these parameters for each previous round. Subject's optimal risk neutral bidding strategy would be to bid half of the valuation, which generates a winning probability of 0.5 (=50%). Therefore, bidding more than half of the valuation represents risk averse behavior, bidding less than half of the valuation stands for risk seeking behavior. The level of risk taking behavior is measured as Constant Relative Risk Aversion (CRRA) according to the index introduced by Arrow and Pratt (1971). The CRRA coefficient is equal to $r = 1 - (\text{valuation} - \text{bid})/\text{bid}$ with an underlying utility function of $U(x) = x^{1-r}$ with $x = \text{expected payoff}$.³ A coefficient of $r < 0$ denotes risk seeking behavior, a coefficient of $r > 0$ means the participant acts risk averse, and the coefficient $r = 0$ represents risk neutral behavior.⁴

²Learning from feedback is explicitly allowed. Thus, participants are offered a table with their endowment, their own bid, the computer's bid (only on output screen), the winning bid, and points won over all previous rounds. To diminish wealth effects, totals are only displayed after the input stage is finished.

³Arrow and Pratt (see Arrow 1971) defined the notion for Constant Absolute Risk Aversion (CARA) and Constant Relative Risk Aversion (CRRA) to measure the utility of risk. Thereafter CARA is measured as $A(x) = -U''(x)/U'(x)$ and CRRA as $xA(x) = R(x) = -xU''(x)/U'(x)$. The utility function $U(x) = x^{1-r}$ used in this study is based on Chakravarty et al. (2011). Solving the function leads to $R(x) = r$, why in this study CRRA is consistent with the utility of the expected payoff and measured as $r = 1 - (\text{valuation} - \text{bid})/\text{bid}$.

⁴For across experiments comparison data need to be aggregated for each individual. Further, the CRRA scale needs to be transformed. Detailed information are given in Sect. 4.3.

Table 1 Risk aversion classification based on lottery choice

Number of low risk choices	Range of relative risk aversion ^a				Risk preference classification	Classification
0–1		r	<	-0.95	Highly risk seeking	Risk seeking
2	-0.95	<	r	<	-0.49	
3	-0.49	<	r	<	-0.15	
4	-0.15	<	r	<	0.15	Risk neutral
5	0.15	<	r	<	0.41	Slightly risk averse
6	0.41	<	r	<	0.68	Risk averse
7	0.68	<	r	<	0.97	Very risk averse
8	0.97	<	r	<	1.37	Highly risk averse
9–10	1.00	<	r			Stay in bed

^a $U(x) = x^{1-r}/(1-r)$, Classification according to Holt and Laury (2002)

3.3 Lottery

As the second experimental vehicle, the Holt and Laury (2002, 2005) compound lottery choice setting is used. The lottery choice treatment consists of ten binary choices between two options that vary with regard to the risk level. The maximum winning amount and the loss amount are higher for the high-risk option (Option B) than for the low risk option (Option A).⁵ The parametrization in points is identical to the Holt and Laury (2002, 1647) experiments (see Table 1). At the end of the experimental session, the winning points are converted into Euro and handed out in cash. The size of the stakes are varied over two separate lottery choice treatments—labelled as “low lottery” and “high lottery” (see Holt and Laury 2002, 2005). Each participant played both lotteries starting with the low lottery. To avoid spillover effects, the information which lottery choice was randomly drawn for actual payoff was revealed only after all choices are made.⁶

⁵The full instructions are presented in the Appendix.

⁶The lottery choice treatment was played in a series of simple experimental games, like ultimatum and dictator game. To avoid spillover effects from these games all entry stages were separated from the outcome stages. After all experiments were played a random number determined which of the experiments was selected for payout. This procedure allows to create a one-shot situation environment. The winning points of the selected experiment were converted into Euro in the end of the laboratory session.

Participants are provided with the information about the winning amounts and the probabilities of winning. These probabilities range from 10% in the first to 100% in the last of the ten decisions. The expected payoff of each binary lottery choice informs the participant about the risk level according to the von Neumann-Morgenstern axioms (1947). Individuals deviating from the risk neutral decision either behave risk seeking or risk averse. Deciding according the expected payoffs signals risk neutral behavior. In the first decision of the low lottery, Option A the expected payoff of $40 \text{ points} * 10\% + 32 \text{ points} * 90\% = 32.8$. Option B of the low lottery reveals an expected payoff of $77 \text{ points} * 10\% + 2 \text{ points} * 90\% = 9.5$ points. Table 1 presents the expected payoffs and the payoff differences between the A and B options in the last three columns. The participants were not provided with this information. According to the expected utility maximization theorem, individuals are assumed to prefer higher over lower expected payoffs. Individuals thus behave risk neutral when choosing Option A in the first four decisions 1 to 4 and then switching to Option B. Individuals behave risk sensitive, when they deviate from this pattern. Earlier switching indicates risk seeking behavior, later switching indicates risk averse decision behavior. The number of Option A choices therefore reveals the level of risk taking preferences. According to Holt and Laury (2002), this measure is translated into CRRA using the utility function for money x of $U(x) = x^{1-r}$, with $r < 0$ indicating risk seeking, $r = 0$ indicating risk neutrality, and $r > 0$ risk averse decision behavior (see Table 1).⁷

3.4 Sample and Experimental Procedure

The participants for the experimental sessions are voluntarily recruited via e-mailing lists at three German universities. A total of 369 students participated in the experiments, 197 of which are female and 172 male. The students are enrolled in different study programs, mainly Business Sciences ($n=178$), Public Administration ($n=119$), and Law ($n=47$). A group of 25 participants is enrolled in different study programs, like Biology, Philosophy, Pedagogy, Gender Studies, or Social Sciences. To control for the impact of the study program, dummy variables are included.

The setup and the user surface of both experimental treatments are programmed and conducted using the software z-tree (Fischbacher 2007). At the beginning

⁷For across experiments comparison data needed to be aggregated for both lotteries. Detailed information are given in Sect. 4.3.

of a session each participant receives a written introduction to the experiment, which is also read aloud.⁸ Both instructions include screenshots to familiarize the participants with the user interface. The participants use randomly assigned computers, which are separated by screens to allow for anonymous playing. After finishing the experimental stages, the participants are presented the survey including socio-demographic characteristics and risk attitudes. Definitions of all variables used in this study are presented in Appendix Table 4.

At the end of a laboratory session, the gained points are converted into Euro using an exchange rate of one point = 0.01€ in the Auction, and an exchange rate of one point = 0.04 € in the Lottery. In total participants, earned between 6.66 € and 22.71 €, including a show-up fee of 3.00 € in the Auction and 4.08 € and 7.70 €, including a 4.00 € show-up fee in the Lottery. Their final payoff was revealed to the participants after all decisions were made and the survey was completed.

4 Empirical Results

4.1 Descriptive and Bivariate Analysis

The experiments are conducted on two different subject samples.⁹ To allow for comparing the outcomes between the two experimental vehicles data are balanced. The balanced dataset is used for the entire analysis. Details on the balancing procedure are given in Sect. 4.3. Appendix Table 6 shows the descriptive statistics separately for both samples, the results of a One-Way Analysis of Variance (ANOVA) of the balanced dataset, as well as total values. As the ANOVA results show, both samples are comparable on the covariates (gender, age, semester, and risk attitudes), after balancing. The Bartlett's statistic informs about whether the ANOVA-assumption of equal variances over the groups is violated. The assumption is fulfilled for all covariates, why both groups can be treated as stemming from one random sample for further analyses. Participants are on average 23.3 years old and 66.4% of them are female.

⁸The full instructions are presented in the Appendix.

⁹To gain statistically reliable results 105 or 51 participants are needed for the experimental or the survey measure respectively. With the given sample size of 369 participants statistical power lies above 99% for the entire analyses. Statistical power is excellent for all analyses except for Model 1 in Table 3, which is below acceptable levels of reliability.

Table 2 Risk attitudes and risk behavior by gender

	Female	Male	Analysis of Variance	Bartlett's test for equal variances
	Mean (SD)	Mean (SD)	Prob>F	Prob> χ^2
Risk attitudes	6.936 (1.718)	5.220 (1.937)	0.000	0.000
<i>Risk behavior</i>				
Both experimental vehicles	0.487 (0.445)	0.304 (0.485)	0.000	0.013
Auction	0.392 (0.405)	0.156 (0.361)	0.002	0.412
Lottery	0.581 (0.464)	0.452 (0.547)	0.052	0.010

Note: Balanced dataset, full sample N=369; Auction: n=119 (Female n=79, Male n=40); Lottery: n=250 (Female n=118, Male n=132), totals are provided in Appendix Table 6

Table 2 shows the self-reported risk attitudes for female and male participants, and the ANOVA and Bartlett's statistics results. Women state significantly higher risk aversion than men. Figure 1 shows the distribution of risk behavior by gender with the CRRA on the x-axis and its density on the y-axis. Over both experimental vehicles females behave more risk averse (mean=0.487, SD=0.445) than males (mean=0.304, SD=0.485) (see Table 2). The ANOVA results show that the level of risk behavior differs significantly between female and male participants.

To control whether the experimental vehicle has an impact on risk behavior by gender, Fig. 2 displays risk behavior separated for female and male participants compared for each of the two experiments. The left hand panel shows that men play less risk averse in the Auction (mean=0.156, SD=0.361) than in the lottery (mean=0.452, SD=0.547). The right hand panel shows that women reveal a similar tendency toward less risk aversion in the Auction (mean=0.392, SD=0.405) than in the Lottery (mean=0.581, SD=0.464). Thus, females and males show slightly different levels of risk behavior under the experimental regimes, but the gender difference appears stable.

Fig. 1 Risk behavior by gender

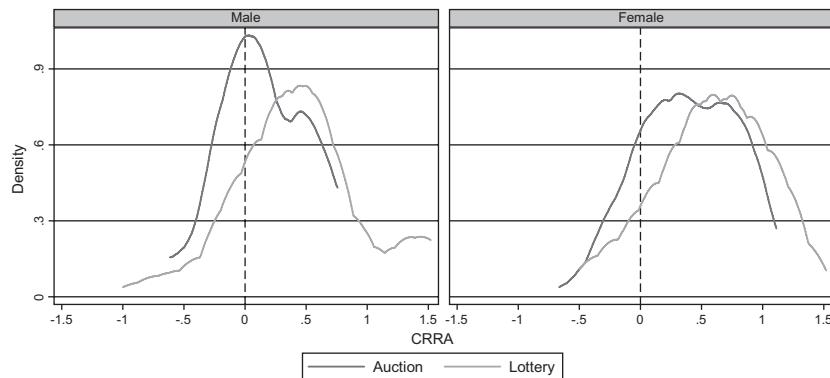
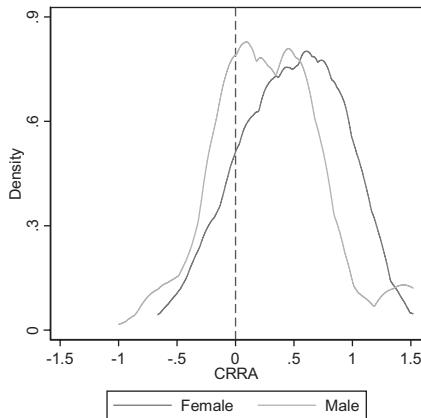


Fig. 2 Risk behavior by experimental vehicle and gender

4.2 Multiple Regression Analysis

Table 3 reports the results of the multiple OLS regression analysis with risk behavior as dependent and the risk attitudes (Model 1 and 3) and gender (Models 2 and 3) as explanatory variables. As the control variables age, semester, the field of study (all Models), and the experimental vehicle (Models 3) are included. The regression results show that self-reported risk attitudes are not likely to predict

Table 3 Determinants of risk behavior

	Model 1	Model 2	Model 3
	CRRA	CRRA	CRRA
Risk attitudes	0.042*** [0.00]		0.020 [0.15]
Female		0.221*** [0.00]	0.179*** [0.00]
Auction			-0.178*** [0.00]
Age	0.011* [0.06]	0.014** [0.02]	0.016*** [0.01]
Semester	-0.014 [0.14]	-0.016* [0.09]	-0.014 [0.15]
<i>Study</i>			
Business	0.06 [0.31]	0.042 [0.45]	0.019 [0.75]
Law	0.428*** [0.00]	0.372*** [0.00]	0.288*** [0.00]
Various	-0.02 [0.81]	-0.068 [0.40]	0.021 [0.80]
Constant	-0.102 [0.55]	-0.022 [0.88]	-0.069 [0.69]
Observations	369	369	369
Adj. R ²	0.076	0.095	0.126

Note: OLS regressions, Study reference category=Public Administration, p-values in brackets

* p < .10, ** p < .05, *** p < .01

risk behavior robustly. When gender is included (Model 3) the previously positive effect of risk attitudes in Model 1 vanishes. Hypothesis H1 therefore needs to be rejected; a direct attitudes-behavior relationship is not supported. In contrast, in line with the descriptive analysis, the results show that women behave significantly more risk averse than men (0.221***, Model 2). This finding remains significant when risk attitudes and the experimental vehicle are included (Model 3), which supports Hypothesis H2, that women behave more risk averse than men. This finding is in line with prior research (Bruhin et al. 2010; Fellner and Maciejovsky 2007; Harrison and Rutström 2008; Holt and Laury 2002, 2005).

The OLS regression results further show that age and risk aversion are positively correlated. This observation is in line with prior research (e.g., Byrnes et al. 1999; Dohmen et al. 2011). Further, the law students among the participants show higher risk averse behavior than the Public Administration students. This indicates that the professional background might impact risk behavior as well. To interpret this finding, however, goes beyond the scope of this study, but can be an interesting starting point for further research.

To elicit the gender effect on the attitudes-behavior relationship, interaction terms between gender and self-reported risk attitudes are included in the full OLS regression model (Table 3, Model 3). The results are presented graphically in Fig. 3.

The left hand panel shows the predicted values of CRRA separated by gender, with the level of self-reported risk attitudes on the x-axis and the predicted values of risk behavior on the y-axis. The graphs show that men and women who report low and medium levels of risk averse attitudes reveal a similar attitudes-behavior relationship. Yet, the relationship between attitudes and behavior fundamentally differs between men and women. For men the flat slope of the graph indicates that the attitudes-behavior relationship is insignificant, i.e. among male individuals, attitudes do not predict behavior. For women there is a positive relationship between attitudes and behavior. Women who report to be more risk averse also tend to behave more risk averse in the experiment.

The middle panel displays the average marginal effects (AME) of gender conditional on the self-reported risk attitudes. The plot demonstrates that the differences between the predicted values for higher levels of self-reported risk aversion are statistically significant, as the confidence intervals do not overlap with zero.

In the right hand panel, the sizes of the AME of risk averse attitudes on risk averse behavior (y-axis) by gender (x-axis) are given. These estimates display the slopes of the predicted values in the left hand panel. Confidence intervals crossing zero indicate that the effect is statistically insignificant. The AME size plot shows that the attitudes-behavior relationship among men is insignificant. The slopes for women are significant at $p < 0.1$.

Overall, these results indicate that the gender gap in risk behavior is moderated by self-reported risk attitudes. Women reporting high levels of risk aversion seem to behave even more risk averse compared to men. For participants who report low levels of risk aversion, gender does not matter. In other words, men behave inconsistent compared to their self-reported attitudes, while for women there is a clear relationship between attitudes and behavior, which means they behave consistently.

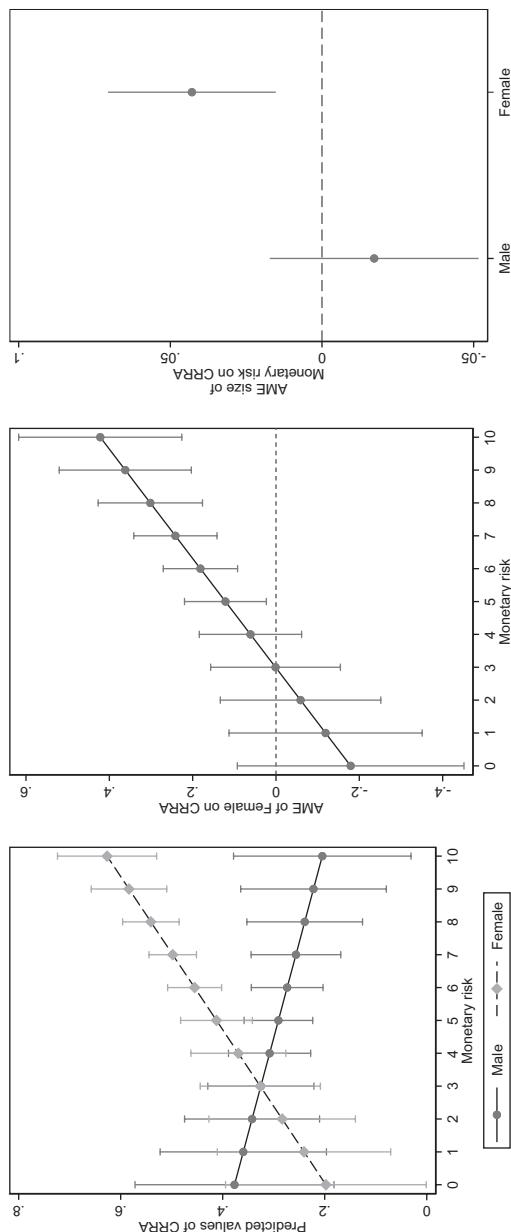


Fig. 3 Gender effect on the attitudes-behavior relationship. Note: Based on full OLS regression model, 90% confidence intervals, from left to right: Predicted values of CRRA by gender; AME of gender on the relationship between risk attitudes and gender; AME sizes of stated risk preferences on risk behavior by gender

4.3 Robustness Checks

First, to compare risk behavior across both experimental vehicles the data structure needs to be adapted. Due to the repeated bidding-procedure, the Auction consists of time series data with ten observations per individual. For the statistical analyses, the observations were aggregated to an average CRRA measure per individual. Appendix Table 9, Model 1 shows the results of a time series regression on the unaggregated data. The period has no impact on the level of risk taking, and the effects of gender and risk attitudes are stable. Findings thus are comparable to the aggregated measures.

Second, the CRRA for the high and the low lottery treatment are aggregated per individual to equalize data structure with the Auction to one observation per subject. Two robustness analyses are presented in Appendix Table 9, Models 2 and 3. Model 2 estimates the full regression model with the CRRA for the high and the low lottery separately. Controlling for the high incentive treatment shows that it is likely to elicit more risk averse behavior than the low lottery treatment (which is consistent with Holt and Laury 2002). The effect of gender and of risk attitudes remains stable. Model 3 shows the regression results for the combined data of the Auction as used in the regular analyses and the observations gathered exclusively from the low lottery treatment. All relevant effects are stable in this regression as well. Thus, using the aggregated lottery data for statistical analyses yields the same results.

Third, the CRRA scale of the original Auction-CRRA needed to be transformed. Originally, Auction data run from -2 (risk seeking)¹⁰ to 1 (risk aversion), but the Holt and Laury (2002) scale runs from -1.17 (risk seeking) to 1.52 (risk aversion). To allow for direct comparison, the Auction scale was transformed to the Holt and Laury scale. Therefore, a CRRA below zero is divided by -2 and multiplied with -1.17 , values above 0 are multiplied with 1.52 . On both scales, a value of zero denotes risk neutral behavior why this point is fixed. Model 5 in Appendix Table 9 shows that all findings remain robust.

Fourth, data are balanced using entropy balancing, which allows to reweight a dataset with a binary treatment so that the distributions (mean, variance, and skewness) of a set of covariates matches between the treatment and the control group (see Hainmueller 2012). On the left-hand sides of Appendix Tables 6 and 8 the descriptive statistics for the unbalanced dataset shows that the two pools reveal equivalent distributions regarding age and most of the self-reported risk

¹⁰After excluding extreme outliers.

attitudes. They differ significantly only with regard to gender, the number of semesters studied, and the level of risk aversion toward other people. The entirety of the statistical analyses are conducted after entropy balancing. As the treatment group the sample of participants in the Auction ($n=119$) is used. Appendix Table 9 Model 4 shows the results of the full OLS regressions for the unbalanced but otherwise transformed data. All findings remain robust. Model 5 reproduces the regression results for the unbalanced and untransformed Auction data and unbalanced Lottery data. These findings also remain robust.

Fifth, the full OLS regression is reproduced employing a CRRA measure that uses continuous values between 1 and 9 according to the number of low risk choices as described in the first column of Table 1. For the Auction the Holt and Laury-scaled CRRA is assigned according to the range of relative risk aversion (see second column of Table 1). Regression results remain robust using these continuous CRRA measures as well.

The sixth and last robustness check addresses the exclusion of four of the six surveyed risk domains. The risk domains car driving, sports and leisure, health, and other people are not considered to predict risk behavior in one of the experiments, because they are not referring to any economic dimension. The results for the two economic risk domains (financial matters and career) are given separately as well. Appendix Table 9, Model 7 shows that the regression results do not change nor that any of the risk attributes predicts risk behavior. Appendix Fig. 1 shows the gender effect on the attitudes-behavior relationship for all six risk attitude measures. The graphs show that none of the non-economic risk attributes has predictive power for risk behavior (see right hand panels) which makes them redundant for this analysis. Thus, the benefit from excluding these four risk attitudes from the analyses for the clear presentation of the results exceeds the loss in completeness, why the exclusion is reasonable. Moreover, the robustness check shows that these risk domains are not related to the economic risk type measured with the Auction and the Lottery which is in line with prior findings (e.g., Dohmen et al. 2011). Both economic risk attitude measures explain the gender effect in the attitudes-behavior relationship as good as the combined risk attitudes variable used in this study.

In conclusion, the robustness checks show that the results reported in the study using data aggregation, scale transformation, balancing, exclusion, and aggregation of risk attributes do not alter any finding reported in the empirical analysis.

5 Conclusion and Discussion

This study analyzes whether self-reported attitudes in economic risk taking predict experimentally revealed risk behavior, and how gender moderates the relationship between both measures. For statistical analysis, data were gathered among 369 students using a survey to elicit self-reported risk attitudes and two monetary incentivized laboratory experiments to measure risk behavior.

The findings can be summarized in three points: First, self-reported risk attitudes are not likely to predict experimental risk behavior directly. Second, in line with prior research (e.g., Bruhin et al. 2010; Harrison and Rutström 2008; Holt and Laury 2002) this study shows that women behave more risk averse than men. Third, this gender gap in risk behavior is moderated by self-reported risk attitudes. Most interestingly, the analyses show that women behave in the experiments consistently to their self-reported risk attitudes, but men do not. The gender difference, however, applies only for individuals who report highly risk averse attitudes. Among men and women who state low levels of risk aversion, gender does not matter for the attitudes-behavior relationship.

Following the theoretic reasoning of Ajzen and Fishbein (1977), the congruence between risk attitudes and risk behavior depends on the congruent perception of the evaluated risk entities. Now, this study finds this doubled congruence only among female individuals who perceive themselves as risk averse. A plausible, yet conjectural explanation for this finding is that risk averse women assess their risk aversion reliably. Men who state risk averse attitudes, in contrast, seem to overestimate their level of risk aversion. Speculating about the drivers of this phenomenon, one could think of a risk related role understanding associating male risk aversion as desirable personality trait. Such self-stereotyping could drive these men to state risk averse attitudes, which consequently find no echo in the outcomes of the experimental task. It is important to note, that the study design ensures a maximally neutral environment deliberately avoiding any framing effects. Therefore, it seems reasonable to conclude that these supposed mechanisms stem from a systematic gender difference in the attitudes-behavior relationship rather than gender specific responses to the elicitation methods (also see Zizzo 2010). However, finding robust empirical evidence for an explanation for the observed gender difference must be left for further research since this goes beyond the scope of this study.

Further, note that the findings of this study are limited to financial risk aversion measured with the employed risk elicitation methods. The strength of this study is its high internal validity through the strictly controlled and monetarily incentivized laboratory environment and the random assignment procedure. However, this comes at the cost of low levels of generalizability. Thus, the interpretation and transfer of the results is recommended only for the used experimental vehicles, survey items, and subject pools. Nonetheless, this reveals a fruitful starting point for research replicating the results using different experimental methods or survey questions or with introducing framing triggering gender specific risk perception, emotions, overconfidence, or real life scenarios. Furthermore, the impact of changing risk preferences through the life cycle might help to understand the observed effects in further depth, testing for instance, whether the attitudes-behavior relationship changes differently among females and males with rising age.

Besides the in-depth analysis in gender risk preferences, this study contributes to the methodological debate in laboratory experimental research. First, both the first-price sealed-bid auction and the Holt and Laury (2002) compound lottery choice treatment provide equivalent measures of financial risk preferences, which is in line with prior findings (e.g., Engel 2011). This finding contributes to the general discourse about the impact of the experimental vehicle on the risk measures. Second, this study contributes to the methodological question whether surveyed and experimentally measured risk preferences are linked, and shows that gender has a moderating impact on the relationship. Both findings help to interpret risk preferences with regard to gender and the elicitation method and should be informative for future research.

Funding

This study received funding from the Fritz Thyssen Foundation (FTF, Az. 10.13.1.113). The FTF is not responsible for any of the findings reported here.

Appendix

Displayed instructions—Auction

In the following auction you will bet against a computerized bidder over 30 rounds. Your goal is to outbid the computer. If your bid is higher than the computer's bid you have won the respective round. In the case of ties your bid

does not win. In each round, you will have a new randomly chosen endowment with an amount between 1 and 100 points. The computer always bids between 1 and 50 points. All amounts are distributed equally.

$$\text{Gain} = \text{endowment} - \text{bid}$$

All the points you win are added, and converted into Euro with an exchange rate of

$$1 \text{ point} = 0.01 \text{ €}.$$

For each bid helpful information is provided (see Fig. 2 of the paper instructions). The following data can be displayed as often as you wish:

Expected payoff: Outcome in points in the case of winning

$$\text{Expected payoff} = \text{endowment} - \text{bid}$$

Winning probability: Probability of winning the auction with the actual bid. The value ranges from 0 to 1. The higher the value, the higher is the probability to win. 0=sure no win, 0.5=same probability of winning or not winning, respectively, 1=sure win

$$\text{Expected value} = \text{expected payoff} * \text{winning probability}$$

Example:

With an endowment of 59 points you bid 27 points, and have a winning probability of 0.5 (i.e. 50%). The computer's bid is 13 points. Your bid is higher than the bid of the computer, thus you are the winner of the auction. In this round you are winning

$$59 - 27 = 32 \text{ points} * 0.01 \text{ €} = 0.32 \text{ €}.$$

As soon as you have decided on your bid, please select the appropriate row and click the button "submit bid".

Displayed instructions—Lottery Structure

In this game you choose between alternative lotteries.

You will decide 10 times. Each time you choose between option A and option B.

Each option can lead to two events. The probability and the gain in points of each event will be displayed in a table.

Please notice that only **one** of your decisions will determine your earnings.

Previously you will not know which lottery will be played.

When the time for input expired the computer will calculate two random numbers.

The first random number (between 1 and 10) determines which lottery will be played.

The second random number (between 1 and 10) determines your gain in points depending on your choice for option A or B.

Depending on the option you chose (A or B) the respective amount in points will be added to your account.

Example

Notice the example on the following screen. Please click “Forward”.

Lottery number	Option A	Option B	Your decision (A or B)
1	40 of 10% or 32 of 90%	77 of 10% or 2 of 90%	
2	40 of 20% or 32 of 80%	77 of 20% or 2 of 80%	
3	40 of 30% or 32 of 70%	77 of 30% or 2 of 70%	A
4	40 of 40% or 32 of 60%	77 of 40% or 2 of 60%	
5	40 of 50% or 32 of 50%	77 of 50% or 2 of 50%	

Example

If the first random number is 3, the lottery number 3 will be played. We assume you chose Option A in lottery number 3. Option A in lottery number 3 provides “40 of 30% or 32 of 70%”. In other words Option A in lottery number 3 offers a 30% chance to win 40 points and a 70% chance to win 32 points. If you would have been chosen Option B in lottery number 3 you had have a 30% chance to win 77 points and a 70% chance to win 2 points. The second random number determines which event actually occurs. Imagine a 10-sided die. If the throw of the 10-sided die in Option A of lottery number 3 is smaller than or equal 3 you win 40 points. If the throw of the 10-sided die in Option A of lottery number 3 is greater than 3 you win 32 points. Choose one of the options in each of the 10 lotteries. There are two runs of this game each with different amounts of points.

Appendix (Tab. 4, 5, 6, 7, 8 und 9; Fig. 4)

Table 4 Lottery choice experiment by Holt and Laury (2002)

Low lottery		Option B										EP(A) EP(B) EP(B) – EP(A)					
No.	Option A	with	10%	or	32	with	90%	77	with	10%	or	2	with	90%	32.8	9.5	-23.3
1	40	with	10%	or	32	with	80%	77	with	20%	or	2	with	80%	33.6	17	-16.6
2	40	with	20%	or	32	with	70%	77	with	30%	or	2	with	70%	34.4	24.5	-9.9
3	40	with	30%	or	32	with	60%	77	with	40%	or	2	with	60%	35.2	32	-3.2
4	40	with	40%	or	32	with	50%	77	with	50%	or	2	with	50%	36	39.5	3.5
5	40	with	50%	or	32	with	40%	77	with	60%	or	2	with	40%	36.8	47	10.2
6	40	with	60%	or	32	with	30%	77	with	70%	or	2	with	30%	37.6	54.5	16.9
7	40	with	70%	or	32	with	20%	77	with	80%	or	2	with	20%	38.4	62	23.6
8	40	with	80%	or	32	with	10%	77	with	90%	or	2	with	10%	39.2	69.5	30.3
9	40	with	90%	or	32	with	0%	77	with	100%	or	2	with	0%	40	77	37
High lottery		Option B										EP(A) EP(B) EP(B) – EP(A)					
No.	Option A	with	10%	or	80	with	90%	192.5	with	10%	or	5	with	90%	82	23.8	-58.2
1	100	with	10%	or	80	with	80%	192.5	with	20%	or	5	with	80%	84	42.5	-41.5
2	100	with	20%	or	80	with	70%	192.5	with	30%	or	5	with	70%	86	61.3	-24.7
3	100	with	30%	or	80	with	60%	192.5	with	40%	or	5	with	60%	88	80	-8
4	100	with	40%	or	80	with	50%	192.5	with	50%	or	5	with	50%	90	98.8	8.8
5	100	with	50%	or	80	with	40%	192.5	with	60%	or	5	with	40%	92	117.5	25.5
												(Fortsetzung)					

Table 4 (Fortsetzung)

Low lottery		Option B						EP(A)		EP(B)		EP(B)		EP(B)		EP(B) – EP(A)		
No.	Option A	100	with	70%	or	80	with	30%	192.5	with	70%	or	5	with	30%	94	136.3	42.3
7	100	with	70%	or	80	with	30%	192.5	with	70%	or	5	with	30%	94	136.3	42.3	
8	100	with	80%	or	80	with	20%	192.5	with	80%	or	5	with	20%	96	155	59	
9	100	with	90%	or	80	with	10%	192.5	with	90%	or	5	with	10%	98	173.8	75.8	
10	100	with	100%	or	80	with	0%	192.5	with	100%	or	5	with	0%	100	192.5	92.5	

Note: EP = expected payoff. The last three columns were not presented to subjects. EP(A) = expected payoff from choosing option A. EP(B) = expected payoff from choosing option B. The grey area indicates the EP maximizing choice which is option A in Lottery 1 to 4 and option B in Lottery 5 to 10

Table 5 Variables

Variable	Item/Definition
<i>Risk behavior</i>	
CRRA	Constant Relative Risk Aversion [-1.17 highly risk seeking, 0 risk neutral, 1.52 highly risk averse] (according to Holt and Laury (2002)) Auction: $r = 1 - (\text{valuation} - \text{bid})/\text{bid}$ (transformed to [-1.17, 1.52])
	Lottery: $U(x) = x^{1-r}/(1-r)$
Auction	Experimental vehicle [1 – Auction, 0 – Lottery]
<i>Risk attitudes</i>	
(Monetary) Risk attitudes	Mean value of risk attitudes towards Financial matters and Career
Financial matters Career Car driving Sports and leisure Health Other people	How do you see yourself: Are you generally a person who is fully prepared to take risk or do you try to avoid taking risk? How do you see your willingness to take risk in the following subsections? [0 Fully prepared to take risk – 10 Highly risk averse] (according to GSOEP 2004, see Dohmen et al. (2011), scale reversed for analysis)
<i>Sociodemographics</i>	
Female	Gender [1 – Female, 0 – Male]
Age	Age in years
Semester	Number of semesters [1–14]
Study	Study program [1 – Public Administration, 2 – Business, 3 – Law, 4 – Various]
Period	Repetitions of Auction [1–10]

Table 6 Descriptive statistics by experimental vehicle and bivariate analysis for the unbalanced and balanced data

(Fortsetzung)

Table 6 (Fortsetzung)

Variable	Unbalanced				Balanced				Total Bartlett's test for equal variances	Analysis of Variance Prob> χ^2	Total Mean (SD)
	Auction	Lottery	Analysis of Variance Prob>F	Total Mean (SD)	Auction	Lottery	Analysis of Variance Mean (SD)	Prob>F			
CRRRA high lottery	0.540							0.598			
	(0.555)								(0.551)		
CRRRA female	0.392 (0.405)	0.550 (0.454)	0.013 (0.004)	0.274 (0.441)	0.487 (0.405)	0.392 (0.464)	0.581 (0.464)	0.003 (0.000)	0.532 (0.150)	0.487 (0.445)	
CRRRA male	0.156 (0.360)	0.419 (0.533)	0.004 (0.487)	0.005 (0.361)	0.358 (0.361)	0.156 (0.547)	0.452 (0.547)	0.000 (0.250)	0.150 (0.250)	0.304 (0.485)	
N	119	250			369	119	250		369		

Table 7 Risk attitudes and risk behavior by gender

	Female	Male	Analysis of Variance	Bartlett's test for equal variances
	Mean (SD)	Mean (SD)	Prob > F	Prob > χ^2
(Monetary) Risk attitudes	6.936 (1.718)	5.220 (1.937)	0.000	0.041
<i>Risk attitudes</i>				
Financial matters	7.769 (1.824)	5.696 (2.569)	0.000	0.341
Career	6.104 (2.195)	4.744 (2.279)	0.000	0.002
Car driving	6.365 (2.644)	5.282 (2.550)	0.000	0.000
Sports and leisure	4.265 (2.260)	3.286 (1.993)	0.000	0.000
Health	6.489 (2.708)	6.177 (2.601)	0.290	0.000
Other people	6.496 (2.235)	5.596 (2.362)	0.000	0.004
	197	172		

Note: Balanced dataset, full sample N = 369; Auction: n = 119 (Female n = 79, Male n = 40); Lottery: n = 250 (Female n = 118, Male n = 132), totals are provided in Appendix Table 8

Table 8 Descriptive statistics by experimental vehicle and bivariate analysis for the unbalanced and balanced data, all risk attitudes

Variable	Unbalanced			Balanced					
	Auction	Lottery	Analysis of Variance	Total	Auction	Lottery	Analysis of Variance	Bartlett's test for equal variances	
Financial matters	7.008	6.724	0.281	0.762	7.001	7.008	7.136	0.597	0.000
	(2.402)	(2.346)			(2.402)	(2.402)	(2.239)		(2.319)
Career	5.663	5.372	0.254	0.987	5.664	5.663	5.629	0.886	0.000
	(2.297)	(2.294)			(2.297)	(2.297)	(2.335)		(2.312)
Car driving	5.882	5.960	0.799	0.634	5.882	5.882	6.120	0.392	0.000
	(2.662)	(2.765)			(2.662)	(2.662)	(2.660)		(2.659)
Sports and leisure	3.958	3.800	0.514	0.228	3.958	3.958	3.914	0.849	0.000
	(2.312)	(2.104)			(2.312)	(2.312)	(2.132)		(2.220)
Health	6.202	6.012	0.527	0.713	6.202	6.202	6.567	0.190	0.000
	(2.638)	(2.716)			(2.638)	(2.638)	(2.705)		(2.673)
Other people	6.193	5.448	0.007	0.217	6.193	6.193	6.193	1.000	0.000
	(2.319)	(2.560)			(2.319)	(2.319)	(2.319)		(2.315)
N	119	250			369	119	250		369

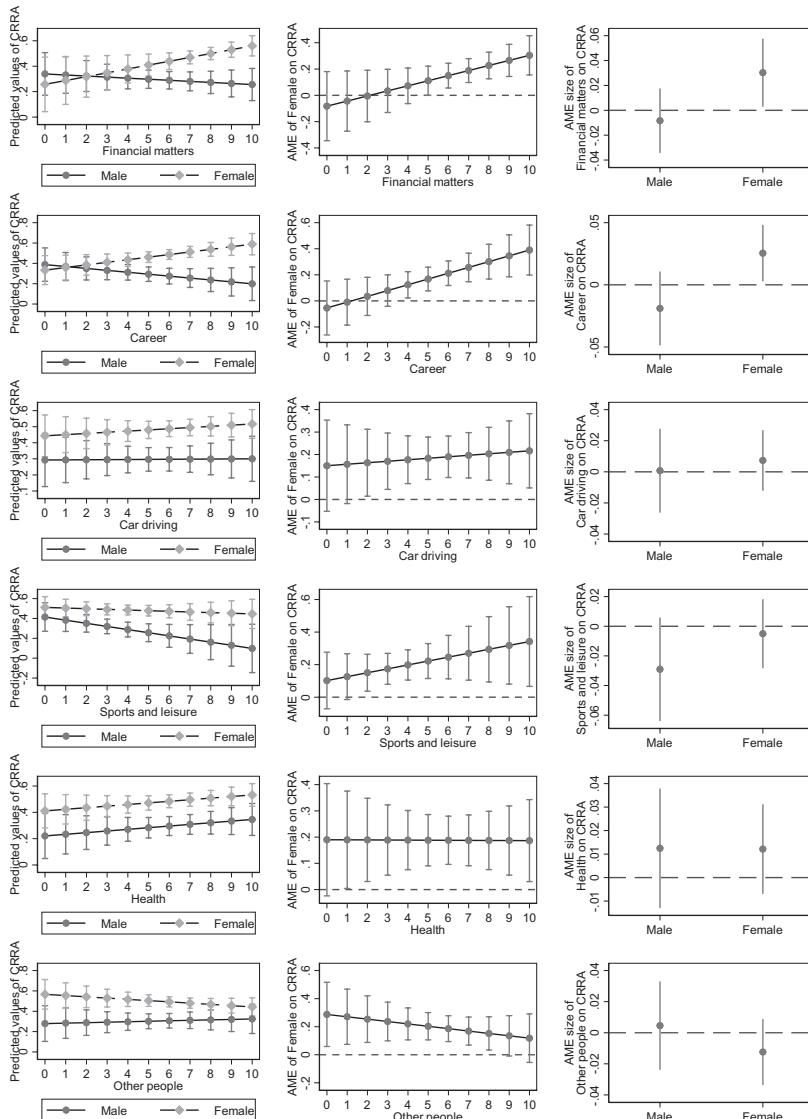


Fig. 4 Gender effect on the attitudes-behavior relationship, all risk attitudes. (Note: Based on full OLS regression model, 90% confidence intervals, from left to right: Predicted values of CRRA by gender; AME of gender on the relationship between risk attitudes and gender; AME sizes of stated risk preferences on risk behavior by gender)

Table 9 Robustness check—Risk behavior

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	Auction	Lottery	Auction + Low lottery	Unbalanced	Continuous	CRRRA	CRRRA
	CRRRA	CRRRA	CRRRA	CRRRA	CRRRA	CRRRA	CRRRA
High Incentive		0.118*** [0.00]					
Period	-0.001 [0.86]						
Female	0.175** [0.02]	0.145** [0.03]	0.181*** [0.00]	0.157*** [0.00]	0.143*** [0.00]	0.642*** [0.00]	0.188*** [0.00]
(Monetary) Risk attitudes	0.020 [0.27]	0.018 [0.31]	0.021 [0.15]	0.021 [0.13]	0.018 [0.16]	0.056 [0.22]	
<i>Risk attitudes</i>							
Financial matters						0.008 [0.50]	
Career						0.012 [0.34]	
Car driving						0.005 [0.60]	
Sports and leisure						-0.014 [0.27]	
Health						0.012 [0.18]	
Other people						-0.006 [0.54]	

(Fortsetzung)

Table 9 (Fortsetzung)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Auction	Auction	Lottery	Auction + Low lottery	Unbalanced			
CRRRA	CRRRA	CRRRA	CRRRA	Raw CRRRA	CRRRA	CRRRA	CRRRA
Auction		-0.109*	-0.182***	-0.306***	-0.790***	-0.175***	
		[0.05]	[0.00]	[0.00]	[0.00]	[0.00]	
Age	0.000	0.018*	0.012*	0.012**	0.042**	0.016**	
	[0.98]	[0.06]	[0.05]	[0.04]	[0.04]	[0.03]	
Semester	-0.018	0.001	-0.01	-0.006	-0.005	-0.044	-0.013
	[0.15]	[0.97]	[0.34]	[0.58]	[0.62]	[0.18]	[0.19]
Study							
Business	-0.035	0.081	0.046	0.056	0.061	0.015	0.018
	[0.69]	[0.25]	[0.47]	[0.35]	[0.29]	[0.94]	[0.76]
Law	0.027	0.118	0.320***	0.134	0.134	0.854***	0.272***
	[0.74]	[0.27]	[0.00]	[0.13]	[0.11]	[0.01]	[0.01]
Various							
	0.035			0.042	0.029	0.172	0.024
				[0.69]	[0.69]	[0.78]	[0.54]
Constant	0.251	-0.234	-0.095	-0.029	-0.01	4.355***	-0.095
	[0.25]	[0.35]	[0.61]	[0.87]	[0.95]	[0.00]	[0.60]
Observations	1155	500	369	369	369	369	369
Adj. R ²		0.040	0.086	0.058	0.116	0.141	0.122

Note: Model 1: random effects OLS time series regression, Model 2: OLS regression clustered by subject, Models 3–7: OLS regression, CRRRA according to Holt and Laury (2002) accept Model 5 (Auction data); Raw CRRRA; r = 1 – (valuation – bid)/bid [–2risk seeking; 1 risk aversion] with 0 being risk neutral, and accept Model 6: CRRRA as continuous variable [1 risk seeking; 0 risk aversion] with 4 being risk neutral. Study reference category = Public Administration, p-values in brackets

* p < .10, ** p < .05, *** p < .01

References

- Ajzen, Icek, and Martin Fishbein. 1977. Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research, *Psychological Bulletin* 84 (5): 888–918.
- Arrow, Kenneth Joseph. 1971. *Essays in the Theory of Risk-Bearing*. Vol. 1. Markham Publishing Company Chicago.
- Berg, Joyce, John Dickhaut, Kevin McCabe, and Vernon L. Smith. 2005. Risk Preference Instability across Institutions: A Dilemma, *Proceedings of the National Academy of Sciences of the United States of America* 102 (11): 4209–14.
- Bohner, Gerd, and Nina Dickel. 2011. Attitudes and Attitude Change, *Annual Review of Psychology* 62: 391–417.
- Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. 2010. Risk and Rationality: Uncovering Heterogeneity in Probability Distortion, *Econometrica* 78 (4): 1375–412.
- Byrnes, James P., David C. Miller, and William D. Schafer. 1999. Gender Differences in Risk Taking: A Meta-Analysis, *Psychological Bulletin* 125 (3): 367.
- Chakravarty, Sujoy, Glenn W. Harrison, Ernan E. Haruvy, and E. Elisabet Rutström. 2011. Are You Risk Averse over Other People's Money?" *Southern Economic Journal* 77 (4): 901–13.
- Charness, Gary, and Uri Gneezy. 2012. Strong Evidence for Gender Differences in Risk Taking, *Journal of Economic Behavior & Organization* 83 (1): 50–58.
- Charness, Gary, Uri Gneezy, and Alex Imas. 2013. Experimental Methods: Eliciting Risk Preferences, *Journal of Economic Behavior & Organization* 87: 43–51.
- Cohn, Alain, Ernst Fehr, and Michel André Maréchal. 2017. Do Professional Norms in the Banking Industry Favor Risk-Taking?, *The Review of Financial Studies* 30 (11): 3801–23.
- Croson, Rachel, and Uri Gneezy. 2009. Gender Differences in Preferences, *Journal of Economic Literature* 47 (2): 448–74.
- Daruvala, Dinky. 2007. Gender, Risk and Stereotypes, *Journal of Risk and Uncertainty* 35 (3): 265–83.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences, *Journal of the European Economic Association*, 9: 522–50.
- Eckel, Catherine C., and Philip J. Grossman. 2002. Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk, *Evolution and Human Behavior* 23 (4): 281–95.
- Eckel, Catherine C., and Philip J. Grossmann. 2008. Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices, *Journal of Economic Behavior & Organization* 68 (1): 1–17.
- Engel, Russell. 2011. First Price Auctions, Lotteries, and Risk Preferences Across Institutions, *Journal of Business & Economics Research* 9 (11): 9–27.
- Fehr-Duda, Helga, Manuele de Gennaro, and Renate Schubert. 2006. Gender, Financial Risk, and Probability Weights, *Theory and Decision* 60 (2–3): 283–313.
- Fellner, Gerlinde, and Boris Maciejovsky. 2007. Risk Attitude and Market Behavior: Evidence from Experimental Asset Markets, *Journal of Economic Psychology* 28 (3): 338–50.

- Finucane, Melissa L., Paul Slovic, Chris K. Mertz, James Flynn, and Theresa A. Satterfield. 2000. Gender, Race, and Perceived Risk: The 'white Male' effect, *Health, Risk & Society* 2 (2): 159–72.
- Fischbacher, Urs. 2007. Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, *Experimental Economics* 10 (2): 171–78.
- Gintis, Herbert. 2000. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Behavior*. Princeton, NJ: Princeton University Press.
- GSOEP. 2004. *German Socio-Economic Panel*. Berlin: German Institute for Economic Research (DIW).
- Hainmueller, Jens. 2012. Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies, *Political Analysis* 20 (1): 25–46.
- Halek, Martin, and Joseph G. Eisenhauer. 2001. Demography of Risk Aversion, *Journal of Risk & Insurance* 68 (1): 1–24.
- Harrison, Glenn W., Eric Johnson, Melayne M. McInnes, and E. Elisabet Rutström. 2005. Temporal Stability of Estimates of Risk Aversion, *Applied Financial Economics Letters* 1 (1): 31–35.
- Harrison, Glenn W., and E. Elisabet Rutström. 2008. Risk Aversion in the Laboratory, *Research in Experimental Economics*, 12: 41–196.
- Holt, Charles A. 2007. *Markets, Games, & Strategic Behavior*. Boston, MA: Pearson Addison Wesley.
- Holt, Charles A., and Susan K. Laury. 2002. Risk Aversion and Incentive Effects, *The American Economic Review* 92 (5): 1644–55.
- Holt, Charles A., and Susan K. Laury. 2005. Risk Aversion and Incentive Effects: New Data without Order Effects, *The American Economic Review* 95 (3): 902–4.
- Kahneman, Daniel. 1992. Reference Points, Anchors, Norms, and Mixed Feelings, *Organizational Behavior and Human Decision Processes* 51: 269–312.
- Kraus, Stephen J. 1995. Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature, *Personality and Social Psychology Bulletin* 21 (1): 58–75.
- Kelley, Matthew R., and Robert J. Lemke. 2015. Gender Differences When Subjective Probabilities Affect Risky Decisions: An Analysis from the Television Game Show Cash Cab, *Theory and Decision*. 78 (1): 153–70.
- Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. New York: Hart, Schaffner, and Marx.
- Lemaster, Philip, and JoNell Strong. 2014. Beyond Mars and Venus: Understanding Gender Differences in Financial Risk Tolerance, *Journal of Economic Psychology* 42: 148–60.
- Levin, Irwin P., Mary A. Snyder, and Daniel P. Chapman. 1988. The Interaction of Experiential and Situational Factors and Gender in a Simulated Risky Decision-Making Task, *The Journal of Psychology* 122 (2): 173–81.
- Levin, Irwin P., Stephanie S. Hart, Joshua A. Weller, and Lyndsay A. Harshman. 2007. Stability of Choices in a Risky Decision-Making Task: A 3-Year Longitudinal Study with Children and Adults, *Journal of Behavioral Decision Making* 20 (3): 241–52.
- Lönnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C. Wichardt. 2015. Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison, *Journal of Economic Behavior & Organization* 119: 254–66.

- Maccoby, Eleanor E., and Nathan Maccoby. 1954. The Interview: A Tool of Social Science, Edited by Gardner Lindzey. *Handbook of Social Psychology* 1: 449–87.
- Meier, Kenneth J., Laurence J. O'Toole Jr, and Holly T. Goerdel. 2006. Management Activity and Program Performance: Gender as Management Capital, *Public Administration Review* 66 (1): 24–36.
- Milgrom, Paul R., and Robert J. Weber. 1982. A Theory of Auctions and Competitive Bidding, *Econometrica* 50 (5): 1089–122.
- Morin, Roger, and A. Fernandez Suarez. 1983. Risk Aversion Revisited, *Journal of Finance* 38 (4): 1201–16.
- O'Brien, Diana Z. 2015. Rising to the Top: Gender, Political Performance, and Party Leadership in Parliamentary Democracies, *American Journal of Political Science* 59 (4): 1022–39.
- Pedroni, Andreas, Renato Frey, Adrian Bruhin, Gilles Dutilh, Ralph Hertwig, and Jörg Rieskamp. 2017. The Risk Elicitation Puzzle, *Nature Human Behaviour* 1 (11): 803–9.
- Powell, Melanie, and David Ansic. 1997. Gender Differences in Risk Behaviour in Financial Decision-Making: An Experimental Analysis, *Journal of Economic Psychology* 18 (6): 605–28.
- Schubert, Renate, Martin Brown, Matthias Gysler, and Hans Wolfgang Brachinger. 1999. Financial Decision-Making: Are Women Really More Risk-Averse?" *The American Economic Review* 89 (2): 381–85.
- Von Neumann, John, and Oskar Morgenstern. 1947. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
- Wärneryd, Karl-Erik. 1996. Risk Attitudes and Risky Behavior, *Journal of Economic Psychology* 17 (6): 749–70.
- Zizzo, Daniel John. 2010. Experimenter Demand Effects in Economic Experiments, *Experimental Economics* 13 (1): 75–98.



Reputation Formation in Bargaining. Partners vs. Strangers in the Ultimatum Game

Jan Sauermann

Abstract

The ultimatum game is certainly one of the most extensively studied games in experimental research, yet little is known about the effects of reputation formation in repeated bilateral bargaining. This study presents a laboratory experiment to examine the role of reputation formation by varying the matching procedure. Under partner matching, players interact with the same opponent repeatedly throughout the whole experiment, which allows players to build up a personal reputation. Under perfect stranger matching, subjects meet only once in the course of the experiment and thus, cannot form a reputation of being tough bargainers. The results show that in particular responder behavior is influenced by the variation of the matching procedure. In the partner matching treatment, responders demand a higher share than in the stranger matching treatment. However, this does not lead to decreased efficiency as proposers offer more and thus avoid costly conflicts. Overall, reputation formation strengthens the influence of fairness considerations.

Keywords

Ultimatum game · Bargaining · Reputation · Matching procedure · Laboratory experiment

J. Sauermann (✉)

Cologne Centre for Comparative Politics, Universität zu Köln, Köln, Germany
E-Mail: jan.sauermann@uni-koeln.de

1 Introduction

Bargaining is at the heart of many political processes. At the international level, for instance, state leaders bargain over trade agreements or peace treaties (e.g. Putnam 1988). Moreover, we know of sometimes complex bargains between different levels of government such as the European Union and its member states or between federal governments and regional entities (e.g. Scharpf 1988), and of course, parties bargain over the formation of governments in coalition negotiations (Laver and Shepsle 1996; Laver and Schofield 1998). While game theoretic models have greatly increased our understanding of bargaining processes in the last decades (e.g. Baron and Ferejohn 1989; Scharpf 1997), we still do not fully understand these sometimes complex interactions. In particular, experimental research has demonstrated that behavior consistently deviates from the predictions of game theoretic models because individuals' rationality is limited (Simon 1985), or they are motivated by a mixture of self-interest and social preferences such as fairness and reciprocity (see Roth 1995; Camerer 2003).

To add further complexity, bargaining in politics is usually not a one-shot interaction, but a repeatedly played dynamic game. Thus, actors in bargaining situations often share a common history, knowing each other from prior interactions. Moreover, actors might also anticipate that they probably will meet again after the current interaction and therefore take the consequences of current events for future encounters into account. Consequently, additional aspects like reputation formation and learning play a very important role in dynamic bargaining games.

While there is a large theoretical and empirical literature on the effects of learning in bargaining games (for example Roth and Erev 1995; Slonim and Roth 1998), we know much less about how reputation formation affects behavior in these settings. This is in stark contrast to other areas such as collective action theory, where reputation formation is of central importance in explanations for cooperative behavior in repeated social dilemma games (e.g. Kreps et al. 1982; Axelrod 1986; Ostrom 1998). This paper aims toward addressing this gap by examining the effects of reputation formation in the ultimatum game—a very simple bilateral bargaining game. In the next section, I will introduce the ultimatum game and discuss a theoretical model by Nowak et al. (2000) which shows that reputation formation leads to distributionally fair outcomes in the ultimatum game. Sect. 3 describes my experimental design to test the model's predictions. Sect. 4 presents my results and Sect. 5 concludes.

2 Reputation Formation in Bilateral Bargaining Games

Probably due to its simplicity, the ultimatum game is the most extensively studied game in experimental research on bargaining. In this bilateral bargaining game, one player, called the proposer, makes a take-it-or-leave-it offer. The other player, the responder, decides whether to accept or reject that offer with rejection resulting in both players earning nothing (Güth et al. 1982). If both players only aim at maximizing their own monetary payoffs, the resulting subgame-perfect Nash-equilibrium of the game is for responders to accept all positive offers and consequently, for proposers to offer the lowest feasible amount. In countless experiments, this prediction has been proven wrong. In one-shot encounters, proposers offer on average 40–50% of the distributional amount. Conflict arises frequently, as responders usually reject offers below 25%.¹

Generous out of equilibrium offers can occur because of two different reasons. For one, proposers might be altruistically motivated and therefore make high offers in the ultimatum game. Secondly, proposers might be afraid that responders reject low offers. In this case, proposers still maximize payoffs and strategically offer substantial amounts in order to avoid rejections. A comparison between the ultimatum game and the dictator game provides the opportunity to disentangle whether proposers make high offers because of altruistic or strategic considerations (or both). In the dictator game, the proposer determines how to split the endowment, and the second player has no influence over the resulting distribution (Kahneman et al. 1986). In a direct comparison between both games, Forsythe et al. (1994) show that dictators make lower offers than proposers in the ultimatum game. However, many dictators still offer positive amounts. Hence, proposers in the ultimatum game are both altruistically motivated and strategically try to avoid rejections. Overall, the strategic element seems to be more important (Camerer 2003, p. 44). Hence, bargaining behavior in the ultimatum game is mostly driven by responders' willingness to reject offers that they deem unfair (see also Fehr and Schmidt 1999; Bolton and Ockenfels 2000).

Given the plethora of ultimatum game experiments, it is rather surprising that only relatively little is known about repeated ultimatum play (Avrahami et al. 2013, p. 47).²

¹For overviews see Roth (1995), Camerer (2003), and Güth and Kocher (2014).

²The effect of repetition has only been examined in studies where bargaining partners are randomly re-matched after every period (stranger matching). A meta-study by Oosterbeek et al. (2004) reveals that repeated play of the ultimatum game increases the proposed offers, but does not affect average rejection rates. In another meta-study, Cooper and Dutcher (2011) find that the rejection rate of high offers decreases over time, while the rejection rate of low offers increases.

In particular, the possibility of reputation formation in repeated interactions has received hardly any attention in experimental research on bargaining. A theoretical paper by Nowak et al. (2000) is an important exception.³ In simulations, they analyze the evolution of distributionally fair outcomes in repeatedly played ultimatum games. They show that bargaining outcomes become equal when players can obtain information about previous encounters. Reputation formation by responders drives the effect. The model shows that if responders become known for accepting low offers, proposers might exploit that in the following periods of the game and reduce their offers. Even though rejecting low offers is costly for responders in the short run, in the end responders profit from building a reputation of only accepting high offers. Hence, the model implies that responders should demand a higher share of the endowment when reputation formation is possible in repeated ultimatum games.

The experimental evidence on reputation formation in repeatedly played ultimatum games is extremely sparse. To the best of my knowledge, a study by Slembeck (1999) provides the only direct test of the prediction that responders demand higher payoffs when reputation formation is possible. In the two treatments of his ultimatum game experiments, Slembeck varies the matching procedure for assembling groups in the 20 bargaining periods of the game. The first treatment prevents reputation formation by re-matching bargaining partners randomly at the beginning of every period without giving them information about their opponents' decisions from prior periods. In the second treatment, Slembeck uses partner matching, and subjects play against the same opponent repeatedly throughout the 20 periods of the experiment. Hence, the treatment offers ideal conditions for reputation formation. Slembeck finds no differences in average offers of proposers between treatments. Responder behavior, however, differs in important ways. Overall, rejection rates are higher under partner matching when reputation formation is possible. While rejection rates are similar for low offers in both treatments, rejection rates in the first treatment decrease as offers increase. In the second treatment using partner matching, rejection rates remain high also for moderate offers and only drop when proposers offer almost half of the endowment. Hence, Slembeck's findings confirm the prediction by Nowak et al. (2000) that responders demand a higher share of payoffs when reputation formation is possible.

While Slembeck's (1999) study provides highly interesting empirical insights on the effects of reputation formation in ultimatum games, it relies on a rather low number of observations of 58 participants in the experiment. Moreover,

³See Abreu and Gul (2000) on reputation formation in a more general Rubinstein (1982) bargaining model.

responders can only accept or reject specific offers from proposers, and proposers rarely make low offers below 30% of the endowment. Hence, the experiment reveals only limited information about responders' reactions towards low offers. In the following, I will present an experimental design that provides further empirical evidence on reputation formation in bilateral ultimatum bargaining and furthermore elicits the complete strategy of responders rather than their decisions conditional upon specific offers by proposers.

3 Experimental Design and Procedures

In the experiment, participants play the ultimatum game repeatedly for four periods. In each period, an endowment of 100 points has to be divided between the proposer and the responder. Proposers offer a division of points, whereas a strategy method is used to elicit responders' behavior. Hence, responders do not simply accept or reject specific offers of proposers, but state their minimum acceptable offer without knowing the actual offer of the proposer. Offers below the minimum acceptable offer are rejected, and in this case, both players earn zero points in the period. If the offer matches or exceeds the minimum acceptable offer, the distribution of points is executed as proposed. Using the strategy method does not change the predicted Nash-Equilibrium of the ultimatum game. Hence, responders should still accept any positive offer and thus state the lowest possible positive minimum acceptable offer, i.e. 1 point.⁴ However, the strategy method reveals more information about responders' behavior, in particular for responders' reactions towards low offers than the sequential decision protocol used in Slembeck's (1999) study.

Subjects are randomly assigned to their roles as proposers and responders at the beginning of the first period and play in the same role throughout the whole experiment. I study two treatments. In the *partner matching treatment* (PMT), players interact with the same opponent repeatedly throughout the whole experiment, which allows players to build up a personal reputation. In the *Perfect stranger matching treatment* (SMT), subjects meet only once. To guarantee that two subjects will not be matched twice in the experiment, the matching algorithm generates cohorts of four proposers and four responders. Members from both

⁴Experiments show that using the strategy method has behavioral effects. According to Güth and Tietz (1990) and Blount and Bazerman (1996), it strengthens fairness considerations. However, Oxoby and McLeish (2004) show that differences between sequential decision protocols and the strategy method are rather small.

cohorts interact once and only once in the four periods of the experiment. At the end of a period, proposers receive feedback about the minimum acceptable offer stated by their matched responders, and likewise, responders are informed about the offer of their matched proposers in the period.

Compared to PMT, SMT offers no incentives for reputation building. In PMT, rejecting a moderately positive offer in an early period of the experiment has two consequences for a responder. On the one hand, the responder forgoes payoffs in the current period, but on the other hand, the responder builds up a reputation as a tough bargainer which might trigger higher offers by proposers in the following periods. Under perfect stranger matching, however, the individual responder has to bear the costs of rejecting offers, whereas only those subjects who will face the current opponent in future periods possibly benefit from increased offers.

The experiment was conducted in the Cologne Laboratory for Economic Research using the software z-Tree (Fischbacher 2007).⁵ Subjects were recruited via email using ORSEE (Greiner 2015). Overall, 100 subjects participated in four sessions of the experiment: 28 subjects in PMT and 72 subjects in SMT. No subject participated in more than one session, and all subjects were students from the University of Cologne, most of them students of economics, management, or related fields. In order to avoid income effects, subjects received 0.08€ per point earned in one period which was drawn randomly from the four periods of the experiment. The experiment took about an hour, and on average, subjects earned €13.40.⁶

4 Results

The main findings of the experiment are reported in Tables 1 and 2.⁷ In line with the prediction by Nowak et al. (2000) that responders should demand a higher share of the endowment when reputation formation is possible, the strongest

⁵Instructions and the z-Tree program can be obtained from the author upon request.

⁶This includes payoffs earned in a second part of the experiment. Subjects were not aware of the rules of the second part when they participated in the first part reported here. In detail, payoffs range between €6.50 and €15.60. Subjects earned on average €10.90 as performance-depended payoffs and a show-up-fee of €2.50. Subjects earned one third of the performance-depended payoffs in the first part of the experiment, which lasted about 15–20 min.

⁷In PMT, matched pairs of proposers and responders form 14 independent observations for the statistical analyses. As described above, at the beginning of every period pairs of proposers and responders in SMT are matched from 8-person cohorts consisting of four proposers and four responders. Each 8-person cohort forms an independent observation in the analyses. Overall, SMT yields 9 independent observations.

Table 1 Observed behavior in the partner matching treatment

Period	1	2	3	4	Mean
Mean offer	47.07	49.57	49.14	49.29	48.77
Mean minimum acceptable offer	46.50	47.00	48.14	48.21	47.46
Rejection rate	28.57% (4/14)	0% (0/14)	0% (0/14)	0% (0/14)	7.14% (4/56)
Points earned by proposers	36.14	50.43	50.86	50.71	47.04
Points earned by responders	35.29	49.57	49.14	49.29	45.82
Equal splits	57.14% (8/14)	85.71% (12/14)	78.57% (11/14)	85.71% (12/14)	76.78% (43/56)

Table 2 Observed behavior in the perfect stranger matching treatment

Period	1	2	3	4	Mean
Mean offer	44.86	44.89	47.22	46.78	45.94
Mean minimum acceptable offer	40.14	38.86	38.42	37.58	38.75
Rejection rate	27.77% (10/36)	27.77% (10/36)	25.00% (9/36)	19.44% (7/36)	25.00% (36/144)
Points earned by proposers	38.08	38.25	38.31	42.11	39.19
Points earned by responders	34.14	33.97	36.69	38.44	35.81
Equal splits	36.11% (13/36)	44.44% (16/36)	55.56% (20/36)	52.78% (19/36)	47.22% (68/144)

treatment effect emerges among responders, who demand significantly more points in PMT than in SMT (47.46 vs. 38.75 points, $p < 0.001$, 2-sided Mann-Whitney U-test). The difference arises instantly in the first period (46.50 vs. 40.14 points, $p = 0.008$, 2-sided Mann-Whitney U-test), which is a strong sign of reputation formation among responders in PMT. In the first period, responders have to decide without any experience from prior periods. Hence, different levels of knowledge about the opponent's behavior cannot cause the observed behavioral differences between treatments. Instead, responders in PMT signal that they will not accept any offers fairly below the equal split, whereas responders in SMT, knowing that they will not face the proposer from period one in a later period again, accept on average about six points less. In the course of the experiment, the difference between treatments increases to about 10 points. Looking at the development of minimum acceptable offers over time, responders increase their demands slightly in PMT from 46.50 points in period 1 to 48.21 points in

period 4 ($p=0.133$, Page's non-parametric test for ordered alternatives). In SMT, minimum acceptable offers decrease from 40.14 points in the first period to 37.58 points in the last period ($p=0.636$, Page's non-parametric test for ordered alternatives). In both cases, changes over time do not reach commonly accepted levels of significance.

In contrast to responders, proposers do not try to build up a reputation of making low offers under partner matching. In fact, average offers are significantly higher in PMT than in SMT (48.77 vs. 45.94 point, $p=0.005$, 2-sided Mann–Whitney U-test). Proposers in PMT offer more from the first period on (47.07 vs. 44.86 points, $p=0.061$, 2-sided Mann–Whitney U-test). Hence, proposers seem to anticipate higher demands by responders and strategically offer more in order to avoid conflicts. The desire to avoid rejections can also explain the fact that proposers increase their offers in the course of the experiment in both treatments.

At least in PMT, players are fairly successful in avoiding costly conflicts. While the rejection rate is 28.57% in the first period, it drops to zero in the following periods. In SMT, however, the rejection rate starts at about the same level, but decreases only slightly in the course of the experiment. Even in the last period, 19.44% of all offers in SMT are rejected. Thus, even though responders demand more in PMT, players reach a significantly higher average level of efficiency in PMT than in SMT (7.14% vs. 25.00% rejections, $p=0.005$, 2-sided Mann–Whitney U-test).

As a direct consequence of the lower rejection rate, subjects earn more points in PMT than in SMT. However, although the occurrence of equal splits differs considerably between treatments (76.78% vs. 47.22%), the relative distribution of points is hardly affected. In both treatments, proposers earn on average slightly more points than responders. In PMT, the difference is 1.22 points, and in SMT, it is 3.38 points.

5 Conclusion

This paper provides novel empirical evidence on the prediction by Nowak et al. (2000) that reputation formation increases fairness in repeated ultimatum bargaining. The main results can be stated as follows. Behavior is driven by responders. If responders know that they will face the same opponent in later periods again, they build up a reputation of accepting only offers that are close to or equal to the equal split of the distributable amount. Proposers anticipate responders' behavior and offer more in order to avoid conflicts.

Summing up, partner matching has two effects in the ultimatum game: First, it strengthens the importance of social preferences in bilateral bargaining. Instead of converging on the subgame-perfect Nash-equilibrium, most pairs split payoffs equally throughout the whole experiment. Second, partner matching increases efficiency by reducing the occurrence of conflicts. Overall, these findings add to experimental data suggesting that partner matching bolsters pro-social behavior. For instance, experiments report increased cooperation in public goods games (e.g., Croson 1996; Botelho et al. 2009) and reduced moral hazard problems in trust games (Bolton et al. 2004) under partner matching.

References

- Abreu, Dilip, and Faruk Gul. 2000. Bargaining and reputation. *Econometrica* 68 (1): 85–117.
- Avrahami, Judith, Werner Güth, Ralph Hertwig, Yaakov Kareev, and Hironori Otsubo. 2013. Learning (not) to yield: An experimental study of evolving ultimatum game behavior. *Journal of Socio-Economics* 47: 47–54.
- Axelrod, Robert. 1986. An evolutionary approach to norms. *American Political Science Review* 80 (4): 1095–111.
- Baron, David P., and John A. Ferejohn. 1989. Bargaining in legislatures. *American Political Science Review* 83 (4): 1181–206.
- Blount, Sally, and Max H. Bazerman. 1996. The inconsistent evaluation of absolute versus comparative payoffs in labor supply and bargaining. *Journal of Economic Behavior & Organization* 30 (2): 227–40.
- Bolton, Gary E., Elena Katok, and Axel Ockenfels. 2004. How effective are electronic reputation mechanisms? An experimental investigation. *Management Science* 50 (11): 1587–602.
- Bolton, Gary E., and Axel Ockenfels. 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90 (1): 166–93.
- Botelho, Anabela, Glenn W. Harrison, Lígia M. Costa Pinto, and Elisabet E. Rutström. 2009. Testing static game theory with dynamic experiments: A case study of public goods. *Games and Economic Behavior* 67 (1): 253–65.
- Camerer, Colin. 2003. *Behavioral Game Theory. Experiments in Strategic Interaction*. Princeton: Princeton University Press.
- Cooper, David J., and E. Glenn Dutcher. 2011. The dynamics of responder behavior in ultimatum games: A meta-study. *Experimental Economics* 14 (4): 519–46.
- Croson, Rachel. 1996. Partners and strangers revisited. *Economics Letters* 53 (1): 25–32.
- Fehr, Ernst, and Klaus M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114 (3): 817–68.
- Fischbacher, Urs. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10 (2): 171–8.
- Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton. 1994. Fairness in simple bargaining experiments. *Games and Economic Behavior* 6 (3): 347–69.
- Greiner, Ben. 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1 (1): 114–25.

- Güth, Werner, and Martin G. Kocher. 2014. More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization* 108: 396–408.
- Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* 3 (4): 367–88.
- Güth, Werner, and Reinhard Tietz. 1990. Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology* 11 (3): 417–49.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1986. Fairness and the assumptions of economics. *Journal of Business* 59 (4, pt. 2): S285–S300.
- Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27 (2): 245–52.
- Laver, Michael, and Norman Schofield. 1998. *Multiparty Government: The Politics of Coalition in Europe*. Ann Arbor: University of Michigan Press.
- Laver, Michael, and Kenneth A. Shepsle. 1996. *Making and Breaking Governments. Cabinets and Legislatures in Parliamentary Democracies*. Cambridge: Cambridge University Press.
- Nowak, Martin A., Karen M. Page, and Karl Sigmund. 2000. Fairness versus reason in the ultimatum game. *Science* 289 (5485): 1773–5.
- Oosterbeek, Hessel, Randolph Sloof, and Gijs van de Kuilen. 2004. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics* 7 (2): 171–88.
- Ostrom, Elinor. 1998. A behavioral approach to the rational choice theory of collective action. *American Political Science Review* 92 (1): 1–22.
- Oxoby, Robert J., and Kendra N. McLeish. 2004. Sequential decision and strategy vector methods in ultimatum bargaining: Evidence on the strength of other-regarding behavior. *Economics Letters* 84 (3): 399–405.
- Putnam, Robert D. 1988. Diplomacy and domestic politics: The logic of two-level games. *International Organization* 42 (3): 427–60.
- Roth, Alvin E. 1995. Bargaining experiments. In *The Handbook of Experimental Economics*, eds John. H. Kagel and Alvin. E. Roth, 253–348. Princeton: Princeton University Press.
- Roth, Alvin E., and Ido Erev. 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior* 8 (1): 164–212.
- Rubinstein, Ariel. 1982. Perfect equilibrium in a bargaining model. *Econometrica* 50 (1): 97–109.
- Scharpf, Fritz W. 1988. The joint-decision trap: Lessons from German federalism and European integration. *Public Administration* 66 (3): 239–78.
- Scharpf, Fritz W. 1997. *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Boulder, CO: Westview Press.
- Simon, Herbert A. 1985. Human nature in politics. The dialogue of psychology with political science. *American Political Science Review* 79 (2): 293–304.
- Slembeck, Tilman. 1999. Reputations and fairness in bargaining. Experimental evidence from a repeated ultimatum game with fixed opponents. *Discussion Paper, University of St.Gallen*.
- Slonim, Robert, and Alvin E. Roth. 1998. Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica* 66 (3): 569–96.

Jahrbuch für Handlungs- und Entscheidungstheorie

Eric Linhart · Marc Debus
Bernhard Kittel *Hrsg.*

Jahrbuch für Handlungs- und Entscheidungstheorie

Band 10

 SpringerVS

Jetzt im Springer-Shop bestellen:
springer.com/978-3-658-16713-4

