Spatial Voting across Electoral Arenas and Policy Dimensions

A Panel Analysis of EU Integration Preferences and Electoral Behavior in the 2013-2014 Bavarian Multi-level Elections

Steffen Zittlau

Mannheim, 2018

Inaugural dissertation submitted in partial fulfillment of the requirements for the degree Doctor of Social Sciences in the Graduate School of Economic and Social Sciences at the University of Mannheim

Dekan: Prof. Dr. Michael Diehl, Universität Mannheim

Gutachter: Prof. Thomas Gschwend, Ph.D., Universität Mannheim

Prof. Dr. Marc Debus, Universität Mannheim

Prof. Dr. Peter Selb, Universität Konstanz

Vorgelegt durch: Steffen Zittlau

Datum der Verteidigung: 14.Mai 2018

Abstract

This dissertation proposes a novel way to consistently model policy-based voting behavior across multiple electoral levels. Building on the multidimensional model of spatial competition, change in electoral turnout and party vote choice across elections may result from voters reweighing different policy dimension at different levels of government. An estimation strategy that implements the spatial model in the panel conditional logit fixed-effects framework and allows for the modeling of non-separable preferences is developed. This framework is brought to bear on the long-standing debate on the role of voter EU integration preferences in explaining differences in voting behavior between national-level and European Parliamentary elections. Leveraging a uniquely suited voter panel survey from the German state of Bavaria, evidence of voters recalibrating their policy priorities across electoral levels is established.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most modern democratic polities are organized in a multi-level structure, where citizens are asked to choose representatives and leaders at different levels of government, be it in mayoral, regional or national or transnational elections. Following normative democratic theory, separate elections at each level of government are necessary to legitimize policy making at the specific level. Elections are not only necessary to represent the particular composition of the electorate at the respective geographical unit, but because policy making at different levels ostensibly concerns different aspects of the political, or political issues - while a local legislative may decide on a local building project, a national legislative may decide over unemployment insurance, for example. As political issues vary across the levels, so might partisan alignment: A citizen might agree more with the policy position of one party on local issues, and with the position of a different party on national issues. Only multi-level elections allow citizens to express these level-specific policy, party or candidate preferences. While this reasoning forms the basis for normative democratic theory, it is not clear whether this is actually the case empirically. Do voters differentiate between the different levels of government by factoring in the specific political issues that are at stake at the different levels?

We know relatively little about how level-specific issues influence level-specific voting behavior because this question is located at a traditional blind spot of political science research. Since its early days, political science has been predominantly interested in explaining inter-individual differences in voting behavior in singular elections. Only a minority of studies seeks to explain changes in intra-individual voting behavior, and many of these studies are typically concerned with multiple elections held at the same level at different points in time, normally the most important elections, presidential or national legislative elections. An analytical perspective that focuses on the consequences of change in the level of government on individual voting behavior has so far only taken academic center stage in the field of European studies. This is because the European Union (EU) and its member states are the prime example of a multi-level polity, and a natural laboratory for the study of multi-level electoral democracy.

On the surface, the story of European democracy is one of success. From a consultatory body of delegates sent out by the national parliaments of the member states, the European Parliament (EP) has developed into a directly elected parliament that exerts considerable influence over how the European Union conducts its affairs. Over the years, the Union has successively updated its institutional design with the intent to increase input legitimacy through democratic elections and by tilting the power distribution within the EU institutional framework towards the directly elected body. Article 10 of the Treaty on the European Union declares that the "functioning of the Union shall be founded on representative democracy" and that "citizens are directly represented at Union level in the European Parliament". This has lead many to conclude that procedural input legitimacy of EU policy-making has largely been realized and that the Union is not "so far [...] from transnational democracy" in its institutional form (Habermas, 2012, ix).

More skeptical voices have focused on the particular challenges that legitimation by democratic elections at the European level faces. It has been argued that the EU lacks a "demos" that shares a "Gemeinschaftsglauben" (Max Weber), which is a necessary prerequisite for the universal acceptance of majority decisions (e.g., Majone, 2005). Moreover, European democracy continues to lack the intermediary structure that helps generate a common political sphere. Media, political parties and societal organizations have a predominantly national frame of reference, and therefore fail to aggregate the heterogeneous social, economic, cultural and political realities across member states. Particularly the lack of a competitive European party system that defines "the alternatives of public policy in such a way that the public can participate in the decision-making process" (Schattschneider, 1960, 138) has been identified as a key obstacle for EP elections to generate substantive input legitimacy. Unlike public policy at the national level, EU policies are to a substantial degree about genuinely European issues, first and foremost the conflict over the degree of European integration, i.e., to which extent national sovereignty is transfered to the supra-national EU level. Rather than formulating competitive EU policy alternatives, the main political parties have locked down competition on the EU integration issue dimension by forming a pro-integration cartel (Katz and Mair, 1995). The missing party competition and complicated EU decision-making processes mean that EU policy making is not transparent to European citizens. No transparency ultimately means that citizens are unable to elect representatives based on the issues that are at stake, or to attribute responsibility for EU policy outputs. The democratic quality of EP elections is therefore at odds with both the populist conception of democratic elections as the representation of people's preferences (Dahl, 1971), and the less demanding liberal conception of holding the elected accountable (Riker, 1982).

These particular problems of organizing democratic elections at the European level have lastingly shaped political theory. After the first EP elections in 1979, Reif and Schmitt (1980) proposed a theoretical framework that understands EP elections as second-order *national* elections. A defining characteristic of EP elec-

tions is its low importance, which makes it subordinate to the first-order election, the general national election. Moreover, according to second-order theory, voting behavior in EP elections is not based on factors that are specific to the European political arena, but is primarily influenced by factors that originate from the national arena. Instead of representing citizen's views on genuinely European affairs, voters use EP elections for base motives: to punish national governments for national policy output. Apart from that, voters simply transfer their nationalized decision-making calculus to the European level. Empirically, the second-order elections model has been very successful, and has henceforth become the predominant theoretical lens through which EP elections, and also all other sub-national elections, are understood and analyzed. Apart from academia, the second-order paradigm has also deeply influenced how EP elections are interpreted by political actors and commentators and perceived by the public.

The second-order elections model takes a somewhat cynical view on the feasibility and merit of not only transnational democracy at the European level, but democratic elections at multiple levels of government in general. According to the theory, EP elections cannot fulfill their intended role of legitimizing EU policy making as voters are seen as lacking the mental capacity, or simply not willing, to differentiate between levels of government, the specific issues that are at stake at the different levels or whom to hold responsible. In its extreme form, the second-order model implies that all elections, apart from the first-order election, do not work as intended. Since citizens do not vote based on the specific issues that are at stake at the respective level of government, but merely on the issues at stake at the national level, there is no democratic benefit derived from holding these elections (other than holding the national government accountable throughout the national electoral cycle). Not much would be lost if these elections were abolished, and lower-level representatives apportioned according to party vote shares in the national legislature. Although this is of course not what second-order theorists suggest to be done, it highlights the pessimistic and arguably anti-democratic implications of the model.

In the last two decades, increasingly vocal academic opposition to the second-order interpretation of EP elections has formed. Opposition has rallied behind the proposition that some of the key premises of second-order theory may no longer be valid. The institutional design of the EU has changed substantially since the first EP election in 1979, the range and impact of EU policies has increased, and the intermediary structure has become more Europeanized. Thereby, EP elections have gained importance and have become more "European", in the sense that genuinely European affairs increasingly structure voting behavior in EP elections. Accordingly, this line of literature has been labeled the "Europe Matters" argument. Its proponents do not challenge that EP elections are by far less important to voters than general national elections. However, they challenge the proposition that EP elections necessarily follow a nationalized logic that crowds out all considerations that are specific to the European arena. Over the years, a considerable body of

research has accumulated that suggests that attitudes towards the EU and voter preferences over the extent of European integration are associated with voting behavior in EP elections: European citizens vote for different parties in EP elections, or abstain from voting altogether, because they want to voice their disagreement with the course of the European project and the positions that the parties they usually vote for in national elections take on European affairs. If the Europe Matters argument was correct, the prospect of transnational democracy may be viewed in a much more optimistic light. By increasingly factoring in the specific issues that EP elections are supposed to be about, European citizens might assure that their views on EU policy are represented and that the elected are held accountable. Of course, additional obstacles still stand in the way of full transnational democracy, such as the absence of an executive that is legitimized by the parliament. However, it would suggest that the necessary prerequisites at the level of the electorate are in place for EP elections to potentially fulfill their intended function of generating (or withdrawing) substantive input legitimacy.

Despite its far-reaching implications for the democratic quality of transnational democracy at the European level, the question whether citizens fine-tune their voting behavior to the European context of the EP election or whether they pursue a one-size-fits-all approach to voting by applying their nationalized first-order voting behavior calculus to EP elections has not been adequately resolved by political science research. I argue that part of the reason is that second-order theory has for a long time been the only game in town, and due to its initial success, has monopolized theoretical development not only in EP election research but in multi-level elections research in general. However, since second-order theory assumes a priori that voters follow a one-size-fits-all, nationalized logic of voting, it has understandably not been interested in developing a positive theory of how level-specific political issues shape level-specific voting behavior. While some notable advances towards a model of level-specific issue voting have been made by the Europe Matters literature that has formed in opposition to the second-order literature, these attempts have been piecemeal. As it stands, political science research lacks a general theoretical model to compare policy-based voting behavior across multiple elections, particularly if these elections take place at different levels of government. This dissertation aims to fill this theoretical gap. I propose a generalization of the multidimensional model of spatial competition to multiple elections taking place at different levels of government. The key feature of the model is that it allows for the flexible modeling of variation in the weights that voters assign to multiple relevant policy dimensions. Variation in weights encapsulates the idea that voters might care more about specific political issues at specific levels of government, which in turn might lead them to alter their voting behavior across elections. The model suggests that in multi-party systems, electoral participation and party vote choice not only vary according to the degree to which the relative weight of policy dimensions varies across the electoral levels, but also depends on the multidimensional policy preferences of individual voters, and the

multidimensional spatial configuration of the parties' policy positions. Although the model is developed particularly with regards to a comparison of voting behavior in FOE and EP elections, it is general in nature and applicable to all settings of multi-level democracy.

Another, very much practical reason why political science research has failed to develop a general theoretical model of individual voting behavior across multi-level elections is the preeminence of cross-sectional data structures in empirical voting behavior research. Traditional voter surveys cover only one election - a theory that explains intra-individual changes across multi-level elections simply was not developed since there was no data. In the recent decade, a number of panel study projects have been launched. However many of these panel studies do not allow to compare electoral behavior vertically, since they exclusively focus on elections held at the national level. This dissertation employs new data from a panel study particularly designed for the study of multi-level elections, the Making Electoral Democracy Work (MEDW) Bavaria Voter Panel study. The panel study covers three legislative elections that were held in the years 2013 and 2014 in the German state of Bavaria - a subnational State, a Federal and an EP election - and provides a unique opportunity for a three-way comparison of individual-level voting behavior across levels of government. Moreover, the case of Bavaria constitutes a close to ideal laboratory to study the role of issues in multi-level election behavior since it allows to keep various potentially intervening factors such as institutional and party-level differences across the elections constant.

The discipline's traditional focus on cross-sectional data structures also means that the existing methodologies of comparing individual voting behavior across elections are woefully underdeveloped when compared to other disciplines such as economics and sociology. Even in applications for which new panel data has become available in recent years, voting behavior research, and large parts of political science in general, has been reluctant to adopt statistical techniques appropriate for the analysis of panel data structures. As is well known, panel data can offer unique opportunities to move beyond purely correlational interpretation of findings and come closer to the identification of causal relationships. However, all too often the discipline continues to stick to established cross-sectional techniques that fail to unlock the full potential of panel data. In the field of voting behavior research, these shortcomings may be at least partially explained by the particular challenges that longitudinal analyses here face. Panel attrition rates, and the high monetary costs of keeping them low, mean that most panel data studies can at best include a small number of elections. Statistically speaking, a low number of panel waves limits analytical leverage considerably. Moreover, the phenomena voting behavior research is concerned with, such as vote choice and electoral participation, are discrete in nature. The non-linear models required to study discrete phenomena are ultimately more challenging than linear models on which much of the development and teaching of statistical techniques has traditionally focused. Accordingly, a second goal of this dissertation is the development of a methodology for the empirical

analysis of multi-level voting behavior based upon panel data. My empirical approach seeks to improve the validity of inferences by controlling for unobserved heterogeneity, and at the same time addresses the particular challenges arising from voter survey panel data. I do so by rediscovering the conditional likelihood approach to panel fixed effect discrete choice models which was already developed in the 1980s, but somehow did not find its way into the statistical toolkit of political scientists. As it turns out, the approach is particularly well suited for the short panel studies students of voting behavior are likely to encounter.

The dissertation seeks to answer four specific subquestions: 1) Do voter preferences on EU integration explain differential turnout in EP elections? 2) Do voters care more about EU integration issues when choosing parties in EP elections than in other elections? 3) Does the vital assumption of non-separability of preferences that the multidimensional model of spatial competition relies on, hold? 4) Can vote switching between national and EP elections be explained by the EU integration preferences of voters?

The dissertation is organized as follows:

Chapter 2 discusses the relevant literature. Firstly, I introduce the second-order elections model (Reif and Schmitt, 1980) and critically reflect on one of its pillars in particular, the transfer hypothesis. Secondly, I discuss the Europe Matters literature, which may be understood primarily as a challenge to the transfer hypothesis. Of particular interest to this dissertation is a strand of this literature that argues based on the spatial theory of politics. In the third part, I address important theoretical gaps in these spatial models. I sketch out a consistent two-dimensional spatial voting model that explains multi-level voting behavior as the consequence of variation in the dimensional weights across electoral arenas. I call this the recalibration hypothesis. I discuss how recalibration can explain differential abstention in EP elections and how it can motivate vote switching across FOE and EP elections.

Chapter 3 critically reflects on the state of methodology in the field of multi-level voting behavior research. After clarifying the key obstacle to causal identification in this field of research, I outline how panel data can help make our inferences more reliable. To account for unobserved heterogeneity, I propose the use of the conditional likelihood approach to fixed effects models for discrete choice data. After discussing the ramification of the proposed research designs, I introduce the specific empirical case, panel data set and operationalizations that this dissertation employs.

Chapter 4 investigates the research question whether differential abstention in EP elections can be explained by the EU integration preferences of voters. Two theoretical arguments are tested: Abstention as a consequence of EU legitimacy attributions, or as a consequence of increased policy alienation due to recalibration. The employed fixed effect research design models unobserved heterogeneity, which allows to eliminate various kinds of biases that plague electoral participation

research. I find evidence that differential abstention in EP elections is likely the consequence of increased policy alienation on EU integration preferences in the recalibrated policy space, rather than the consequence of Euroskepticism per se.

Chapter 5 outlines how the recalibration hypothesis can be precisely formulated and directly tested in the multidimensional spatial voting framework. In a Monte Carlo study, I show that a reliable answer will ultimately require taking unobserved heterogeneity into account. This is accomplished by transposing the conditional likelihood approach to a model of multi-party vote choice, the first of its kind in political science research. The empirical evidence is largely consistent with voters assigning a larger relative weight to policy distances on EU integration preferences in EP elections than in FOE or other second-order elections, rather than voters assigning the same relative weight in all elections as implied by the transfer hypothesis.

Chapter 6 primarily consists of a published article (Stoetzer and Zittlau, 2015), that tests a far-reaching assumption commonly made in multidimensional models of spatial competition - that preferences on multiple dimensions are separable. The statistical consequences of violations of the separability assumption are discussed, and a statistical model for the consistent estimation of non-separability is developed. In the second part of the Chapter, I implement the non-separable model specification in the conditional likelihood framework. My findings indicate that voter preferences on Left-Right and EU integration issues are separable, strengthening confidence in the results obtained in Chapter 5.

Chapter 7 investigates the consequences of the larger emphasis that voters put on EU integration issues for differential vote choice at EP elections. The two-dimensional spatial model of voting predicts that whether voters switch support to a different party in EP elections than in FOE is a function of their ideal point on EU integration relative to the party's position they previously voted for, and the position that available party alternatives take on EU integration. I model vote choice at the FOE and EP election as a Markov process, and investigate the transition probabilities in a simulation study. I find that the complex patterns of vote switching predicted by my simple spatial model explain the patterns observed in the data surprisingly well.

Chapter 8 summarizes my findings and highlights the theoretical, methodological and empirical contributions of this dissertation. The implications for future research into multi-level voting behavior and voting behavior research in general are discussed.

Chapter 2

Theories of multi-level voting behavior

2.1 The second-order national elections model and its critiques

After the first European parliamentary (EP) elections in 1979, Reif and Schmitt (1980) proposed an analytical framework to interpret the election results that has strongly influenced the academic understanding of, and public discourse about EP elections ever since. Their theory analyzes EP elections not in their own right, but by setting them in contrast to national general elections (or in presidential systems, the national presidential election), to which Reif and Schmitt refer as first-order elections (FOE). First-order elections are the "decisive elections", in which national governments are elected (p. 8). Elections which do not serve the function of electing the national government, and a national Head of Government, are second-order national elections (SOE). In his later work, Reif (1997, 117) further clarified the term: "All elections (except the one that fills the most important political office of the entire system and therefore is the FOE) are national second order elections, irrespective of whether they take place in the entire, or only in a part of, the country." The term therefore includes all subnational elections such as local, regional and state elections, as well as EP elections. These SOE are characterized by their subordinate role to the first-order electoral arena.

The second-order argument is fundamentally built on saliency. It rests on the observation that there is less at stake in SOE compared to FOE due to the preeminent policy making power of the national government. As Reif and Schmitt (1980, 12) point out, there is even less at stake in EP elections since there is no "European government that is elected, which decreases the relative importance of EP elections even more than in other second-order elections, which at least produce a Head of Government at the local or regional level." Therefore Reif (1997, 121) argued that EP elections might even be called third-order elections. But as this terminology is

not coherently established in the literature, and rather obfuscates the important theoretical points, I shall not use the term throughout this dissertation.

Turnout   A number of empirical predictions about the election results in SOE, relative to FOE, follow directly from the less-at-stake proposition. First of all, SOE turnout is substantially lower than FOE turnout. The primary mechanism is that since there is less at stake, fewer voters will participate in the SOE than in the FOE due to the decreased benefits of voting. Moreover, the lack of politicization of the SOE and lacking mobilization efforts by the parties further depress turnout. The prediction is unequivocally supported by analyses of turnout rates throughout EU member states (except those with compulsory voting, or where FOE and EP elections are held concurrently) (Reif and Schmitt, 1980; Schmitt, 2005; Schmitt and Mannheimer, 1991; van der Eijk and Schmitt, 2008). It is also supported by analyses based on individual-level observational data. Franklin (2004) and Franklin and Hobolt (2011) show that the likelihood to turn out is strongly affected by how voters perceive the relative importance of EP elections. Schmitt (2005) finds the same pattern for the 2004 EP election. The most methodologically persuasive evidence comes from an interesting case study by Orford et al. (2009), who show that perceived costs of voting, like travel distance to the nearest polling station, have a larger effect in lower saliency elections such as EP and local elections, compared to national general elections. A modification to the less-at-stake mechanism that focuses not on individual perceptions, but on the role of party mobilization and media attention has been proposed (Blondel et al., 1997). Here it is the "agencies of mobilization" (Marsh and Mikhaylov, 2010, 11) that perceive to be less at stake in EP elections and devote fewer resources to these campaigns. However, this debate is only over the causal chain that links lower importance to lower turnout, and not the more general mechanism originally proposed.

Winners and losers   Next to the less-at-stake implication of lower turnout in SOE, second-order theory also predicts that small parties do better in SOE than in FOE. The mechanism proposed by Reif and Schmitt (1980) centers on the incentives to vote sincerely or strategically. In FOE, voters that prefer a smaller party often choose to cast their vote for a larger party that is more likely to participate in government formation, or in order to avoid wasting their vote, i.e., they vote strategically. In EP elections, the incentives to vote strategically are diminished. As voters perceive that there is less at stake in EP elections, wasting one's vote is less costly, and since there is no government to be elected, government formation aspects are irrelevant. Therefore more sincere voting is to be expected in SOE, of which small parties are the chief beneficiary. What is disregarded by the classical argument is the role of electoral institutions, which in many member states vary substantially between FOE and SOE. EP elections are oftentimes more proportional, and higher thresholds tend to disadvantage smaller parties in FOE (Kousser, 2004; Prosser, 2016). Accordingly, the relevance of the sincere vs. strate-

gic mechanism should be thought to vary dependent on the degree to which FOE electoral institutions generate incentives to vote strategically and the degree that these incentives are absent in SOE.

A second implication about the winners and losers in EP elections is that government parties tend to do worse than in FOE. Apart from not benefiting from strategic votes, there are two additional mechanisms. Firstly, a voter might simply become disenchanted with the government parties for which he has cast a sincere vote in the FOE due to its performance, and vote sincerely for a different party in SOE (or abstain from voting in SOE altogether). Secondly, a voter might use SOE to signal discontent to the parties in power by temporarily withdrawing support. In such a "protest vote", the desertion is insincere in the sense that the voter remains a long-term supporter of the party, and returns to the party in the following FOE. Protest can also be voiced by strategic abstention of government supporters in SOE, which also increases the vote shares of opposition parties. The temporary withdrawal of support has been also referred to a as instrumental voting (Heath et al., 1999) or voting with the boot (Oppenhuis et al., 1996). Substantively speaking, the SOE takes the character of a referendum on the government rather than being an election in its own right, if the protest motive becomes predominant. These mechanisms indicate that government parties loose support in SOE, dependent on government popularity at the time of the SOE.

An integral part of Reif & Schmitt's framework is that government popularity, and therefore party losses, follow a cyclical pattern. Taking their cues from the US surge-and-decline literature on midterm-elections in the US (Campbell et al., 1960; Tufte, 1975) and similar studies of polling data in the UK (Miller and Mackie, 1973), or the notion of a political business cycle (Nordhaus, 1975), they see government popularity cycles arise somewhat mechanistically. After a short increase of popularity during a "honeymoon period" right after taking office, popularity declines. Popularity reaches its minimum around the midpoint of the electoral cycle, followed by a leveling-out or resurgence as parties ready for the next FOE (see also Marsh, 1998). The magnitude of government party losses then is a function of at which point in time in the national electoral cycle the EP elections are held. While the mechanism that generates the popularity cycle remains debated (Marsh and Mikhaylov, 2010, 12; Weber, 2007), the cyclical pattern of government party losses has been confirmed empirically. The most extensive and convincing empirical investigation of the "winners and losers" argument is Hix and Marsh (2007) and, in a follow-up article, Hix and Marsh (2011). They find strong evidence that smaller parties tend to do better, and large parties to do worse in EP elections than in previous FOE. Government parties tend to do even worse, dependent on the point in the electoral cycle at which EP elections take place.

The transfer hypothesis   In his 1997 clarification, Reif (1997, 117) highlighted the adjective "national" in the original formulation of the second-order *national* elections argument. The term national represents the reasoning that "the campaign

and results of each and every type of SOE are more or less influenced by the political constellation of the dominant political arena within the system, the first order political arena". Although Reif does not rule out that factors specific to the second-order electoral arena might play a role, he suggests that these arena-specific factors fade into the background. The substantive content of SOE, the "campaign and results are more or less heavily influenced by the political constellation of the dominant political arena within the system, the first order political arena (FOPA)" (Reif, 1997, 117). This means that the only defining feature of SOE is that there is less-at-stake. Apart from that, the election is just a copy of the FOE.

Clark and Rohrschneider (2009) call this the *transfer hypothesis*, a term which I shall use throughout this dissertation. Putting it in the perspective of the individual voter, the transfer hypothesis suggests that voters "apply their evaluations of national-level phenomena to the EU level when voting in EU elections" (Clark and Rohrschneider, 2009, 645), or for that matter to all kinds of SOE. In practical terms, this means that when voters make up their mind how to vote in SOE, they think about the *national* economy, the *national* government and the issues that are relevant at the *national* level, and not the politics and policies that are at stake in the SOE.

This perspective on multi-level democracy is provocative and has normative implications. If voters treat all SOE only as less important *national* elections, what's the use of holding these elections? And are the representatives elected in SOE fully democratically legitimized? The interesting and puzzling implications of the transfer hypothesis have contributed much to the popularization of the second-order model and its reception in public discourse. Although maybe not intended by Reif & Schmitt from the outset, the transfer hypothesis has become to be viewed as the central component of the second-order model. In their extensive review of the literature on EP elections, Marsh and Mikhaylov (2010, 13) suggest that the transfer hypothesis, meaning that the issues remain the same in the first- and second-order arena, is indeed "the essence of second-order interpretation".

The insufficient micro-foundations of SOE theory Since the early days of SOE theory, its aggregate-level predictions have been broadly confirmed empirically. Indeed it might be argued that its very success has stifled further theoretical development and critical reflection. Only in the recent decade, mainly coinciding with forthcoming individual-level voter survey data on EP elections, have scholars voiced concerns over the theoretical micro-foundations upon which SOE theory rests. Marsh and Mikhaylov (2010, 10) in their review of the SOE literature point out that "Reif and Schmitt do not offer an explicit theory of a European voter. Their work is essentially an aggregate-level one." In the same vein, Hobolt and Wittrock (2011) have criticized the lack of theoretical micro-foundations and empirical micro-level evidence, that leave ample room for alternative explanations and may invite ecological fallacies. In response to Marsh and Mikhaylov's critique, Schmitt et al. (2008) have defended the mechanisms developed in the original arti-

cle. While Schmitt et al. have helped to clarify the proposed mechanisms, I doubt their defense addresses the fundamental points of the critique. Second-order theory centrally relies on an aggregate-level concept - that there is less at stake in SOE. In a causal language, the difference in importance between FOE and SOE is considered to be the treatment that leads to different outcomes across the elections.[1] That lower importance causes voters to be less likely to turn out is hardly controversial. However, for voters who still turn out in both elections, second-order theory can make conflicting predictions. On the one hand, voters are said to be more likely to vote sincerely in SOE, but on the other hand are expected to vote more strategically, i.e, insincerely. SOE theory does not make a prediction which option prevails in such a situation. This was noted by Hobolt and Wittrock (2011, 31), who laconically point out that "a voter cannot simultaneously vote sincerely *and* strategically". Moreover, according to SOE theory, both options should be understood as dynamic in the sense that which of the two is more relevant depends on the point in time in the electoral cycle where the SOE is held (Reif, 1997). These aspects have not been systematically explored and can lead to conflicting predictions (Marsh and Mikhaylov, 2010, 13).

It is evident that the theoretical micro-foundations of classical SOE are not as solid as they should be. Though it offers a collection of mechanisms, it remains unclear how the different mechanism interact. To be fair, Reif and Schmitt never claimed that the framework developed in their original contribution encompasses a complete theory of individual voting behavior. SOE theory at its outset rather attempted to put the aggregate-level results of EP elections into a coherent framework. The specified mechanisms were seen as giving rise to the aggregate-level patterns. For this purpose, they worked well in their "auxiliary role" to explain macro-level phenomena. But the mechanisms can hardly form the theoretical basis for a model of individual voting behavior in multi-level elections.

A case of methodological nationalism? In a stimulating historiographical review, Schakel and Jeffery (2013) trace back the origins of SOE theory. As Reif and Schmitt readily admitted, the general idea of subordination in importance between elections originates from the work of Reiner Dinkel (1977), who analyzed German Länder elections and found disproportionate losses for parties who were part of the federal government coalition, and that these losses were highest in Land elections that took place around the midpoint of the federal election cycle (the so called "Dinkel-Kurve"). Dinkel (1977, 348) also already stated that Land elections are "subordinate elections [. . .] systematically influenced by the superordinate constellation in the Bundestag". In turn, the origins of Dinkel's reasoning can be found in the "surge and decline" literature on midterm-elections in the US (Campbell et al., 1960; Tufte, 1975) and similar studies in the UK (Miller and

---

[1]Clearly, individuals might vary in their perception of importance differentials, which then varies the treatment intensity.

Mackie, 1973).

Schakel & Jeffery note that the forefathers of the SOE argument had real national elections in mind, in the sense that these elections followed a "nationalized logic of voting behavior" (p. 326). Indeed, midterm elections are clearly real national elections, and Länder elections are also linked to the national arena via the federal state system (Bundesrat). However, regional or European elections do not have this institutional or electoral connection to the first-order national arena. Therefore, the transfer of theory from SOE in which such a electoral connection to the first-order arena is present, to SOE without such a connection may not be valid. From this perspective, it is not a foregone conclusion that *all* SOE are alike and take on a national character. Schakel & Jefferey do not contest that there is less at stake in SOE. However, "what appears less credible is that what is at stake nationally necessarily crowds out distinct judgments about the issues that might be at stake in regional elections." (Schakel and Jeffery, 2013, 327). Although Schakel & Jefferey make their point primarily with respect to regional elections, their argument points to an interesting analytical distinction which has remained rather opaque in the SOE literature: is "how much is a stake" and "what is at stake" the same question? In other words, what is the relation between the concept of second-orderedness, i.e., difference in importance, and the concept of nationalness, i.e. the degree to which the "situation", to use Reif's words (1985, 8), in the second-order arena mirrors the situation in the first-order arena"?

Taking a step back, Reif & Schmitt's original contribution wasn't primarily intended to focus on the question of the nationalization of the European arena. They have never stated that SOE are an exact mirror image of the first-order arena. In his 1997 look-back at the original article, Reif (1997, 16) noted that "[t]he most crucial mistake in the 1980 article was not to have elaborated in sufficient detail how the adjective national in the term 'second order national elections' was meant to be understood." Reif goes on to clarify that "SOE are more or less heavily influenced by the political constellation of the dominant political arena within the system, the first order political arena (FOE)." However, he does not preclude per se the possibility that "arena-specific factors" play a role in determining the outcome of elections in the second-order arena. "[T]he extent to which arena specific factors determine campaign and outcome of SOE varies with inter alia with[sic!] the relative importance attributed by citizens, parties and media and with the degree of nationalization of politics". Reif refers to Caramani (1996; see also Caramani, 2004), who analytically differentiates between two processes of nationalization: horizontal and vertical. While horizontal nationalization means the process of increasing territorial homogeneity, vertical nationalization refers to the process of increasing orientation of voters away from the local and subnational level towards the national level, the first-order political arena.[2] Although Reif has

---

[2]The literature on nationalization has so far not been adequately integrated into the SOE literature. Caramani (2011, 3) points out, while institutional and cleavages aspects of Europeanization have been widely studied, "the formation of a European-wide electorate and party

clearly alluded to (vertical) nationalization, the degree to which voters use the first-order arena as their frame of reference to make up their minds in SOE, in his later writing, the original argument is built on the proposition that all SOE follow a nationalized logic. After all, it is right there in the title of the original contribution - "Nine second-order *national* elections". Moreover, the nationalized logic of SOE has over the years become the central tenet of second-order theorists. Although the process of European integration has advanced since the early days of second-order theory, the transfer hypothesis is maintained by adherents of second-order theory.

Using the terminology of the Lakatosian research program (Lakatos, 1978), the question is whether the transfer hypothesis belongs to the hard core of theoretical assumptions upon which second-order theory rests, or constitutes an auxiliary hypothesis that can be modified. I would argue that the transfer hypothesis should be treated as the auxiliary hypothesis that it is. Originally, it was introduced as a ceteris paribus assumption that allowed the model-builders to focus on the consequences of the difference in importance between the elections, and to ignore other aspects. This was without doubt a legitimate research strategy. However, in the wake of the overwhelming empirical and academic success of the second-order model, the transfer hypothesis became the hallmark of second-order theory. Its counterintuitive and somewhat pessimistic view on voter motives in EP elections - that the issues remain the same even though the EP elections are about something else - contributed much to its intellectual appeal among academia and political commentators. What may also have helped is that the narrative aligns with the preferences of political parties and national governments that seek to avoid a politicization of EP elections (Hix, 1999).

It seems that second-order theorists have unnecessarily tied themselves down to the transfer hypothesis. On theoretical grounds, the transfer hypothesis can not be upheld. In the absence of a direct institutional electoral link to the first-order arena, it simply does not follow that 'what is at stake' is crowded out by that there is 'less at stake' in EP elections.

2.2   The Europe Matters perspective

Over the years, an alternative strand of research has developed that has been labeled the "Europe Matters" argument (Hix and Marsh, 2007). The research

---

system remains an under-researched aspect". Indeed, this strand of research could potentially provide additional theoretical and analytic leverage for the study of voting behavior in European and national elections, but still fails to be integrated in the wider research of EP elections and their relation to national elections. Caramani (2011, 3) argues that the dominance of SOE theory is partly responsible for that, since it "addressed the national character of European elections", leading the focus of research away from the "commonalities in behavior throughout Europe". An exception is Hix and Lord (1997), who have transposed theories of nationalization to describe and explain the process of Europeanization.

program commonly subsumed under this label is highly unstructured and patchy. Many contributions speak only about specific phenomena, party families, member states or EP elections. Furthermore, there is no common theoretical framework that guides the research program. What the contributions however all have in common is that they suggest that "European" factors increasingly play a role in explaining voting behavior. This proposition is seen as being at odds with the second-order model's transfer hypothesis that sees all SOEs following a nationalized logic. Europe matters scholars do not reject the less-at-stake explanation of second-order theory. What they reject is the singularity of the less-at-stake mechanism as the only mechanism that generates different voting behavior in EP elections. So rather than proposing an alternative approach, the Europe matters argument proposes a modification of the transfer hypothesis, and an extension of the second-order model by a second mechanism. The second mechanism they advocate for focuses on factors that are specific to the EP election policy arena. These factors are conceptually independent of the less-at-stake considerations of voters. The common analytical starting point of the Europe Matters argument is the institutional and political context of EP elections, which has substantially changed since the first EP election in 1979. EU treaty amendments have successively increased the powers of the European Parliament and the economic, institutional and political integration of the European Union. Proponents of the Europe Matters perspective argue that this has led to an increase in the importance of EP elections relative to FOE, which has altered the fundamentals upon which the early second-order model was built. Moreover, political discourse, although largely confined to the elite level, and changes in EU member state political opinion and party systems have successively integrated European aspects. Therefore EP elections are said to not only have become more important, but to increasingly take on European character. The European character expresses itself as such that factors that are specific to the European policy arena play a role in explaining EP election voting behavior.

The European policy space   The Europe Matters argument is to be seen against a broader analytical background. On the level of party competition, political conflict in the national arena of EU member states is primarily structured by a general left-right dimension. The left-right dimension is the ideological super-issue for parties and voters (Downs, 1957). While its concrete meaning may vary between countries and over time (Benoit and Laver, 2006; de Vries and Marks, 2012), its defining property is that a large number of specific policy issues are subsumed under this super-issue (Pierce, 1999, 30). As is generally accepted, the major parties are traditionally geared towards, and indeed built for, competition along this left-right dimension (Lipset and Rokkan, 1967). However, the process of European integration gave rise to a new dimension of political conflict. EU policy making is not only structured by the traditional left-right dimension, but contains a strong element of competition over the degree of integration. That is the degree

to which policy making powers are transferred from the national level to the European level, or vice versa. This new dimension that captures the conflict over more national independence versus more integration is generally referred to as the national sovereignty or European Integration dimension. On a tangent, the Europe Matters argument here implicitly refers to models of a European policy space that have been proposed in the literature (see for an overview Steenbergen and Marks, 2004). The Hix-Lord model (Hix and Lord, 1997) sees EU policy preferences as structured by a left-right dimension and a national sovereignty dimension. These dimensions are orthogonal to each other, as parties are unable to incorporate political conflict on EU policy in the left-right, domestic dimension of contestation. The major parties and social classes are internally split, since the question of national sovereignty cuts across the cleavage between functional groups that describes the left-right dimension of domestic contestation. The Hooghe-Marks model (Hooghe and Marks, 2001) argues that parties are only to some degree able to integrate EU policy issues into the contestation along the left-right dimension. Some aspects of EU policy can be absorbed into the left-right dimension, while other aspects which touch conflict over nationalism or supranationalism remain distinct from the left-right dimension. This leads to a configuration of the political space in which the two dimensions remain distinguishable, but are not orthogonal to each other.[3] Party competition tends to be structured by two political camps, Regulated capitalism and Neoliberalism. Center-left parties tend to favor increased European integration for the sake of economic market regulation, while the political right wants to limit regulation of markets on the European level.[4]

---

[3]Ray (2004) proposes a variant of the Hooghe-Marks model that allows for variation of the policy space across EU countries. Following Ray, the way in which left-right and EU integration dimension are related to each other depends on policy expectations. The basic idea is quite simple and convincing: If parties and citizens can expect the EU level to implement a policy that is closer to their ideal point than the policy expectation at the national level, these forces should support the transfer of policy decision-making power to the European level, i.e. be in favor of deeper integration of the European Union. In turn, social groups that can expect better policy returns if policy is implemented on the national level should take on Europskeptic positions. Ray's contribution therefore suggests that the configuration of the policy space varies between countries, to the degree that policy expectations also vary across countries.

[4]The two-dimensional model is not uncontested. The international relations (IR) model (Steenbergen and Marks, 2004) sees political conflict about EU policy as structured along a single dimension that ranges from preferences for less integration to preferences for more integration. This dimension is independent of domestic concerns, that are structured along a traditional left-right dimension. The IR model regards preferences on EU policy from a purely foreign policy perspective, and focuses o the preferences of national governments in pursuit of national interests, not so much the preferences of domestic audiences, i.e. the European citizens or political parties. Alternatively, the so called regulation model of Tsebelis and Garrett (2000) sees preferences on EU policy-making as structured vis-a-vis the extent of economic regulation in EU policy. This corresponds to the domestic left-right dimension, with leftist parties pursuing more economic regulation across Europe, and rightist parties favoring less regulation. Instead of remaining separate, preferences over national sovereignty are successively subsumed into the domestic dimension of conflict - left-right. Given the two-dimensional structure of the European policy space, the general consensus is that traditional parties follow their natural inclination

European issues   The early years of European integration were marked by a pro-integration consensus among centrist parties, particularly in the economic sphere. This "permissive consensus" has also been pointedly described, following the analogy of the political market, as a cartel (Sitter, 2001, 36). By forming such a cartel, the major parties effectively eliminated competition on the European integration dimension. However it also meant that the supply side, the parties, failed to accommodate for the demands of centrist voters that oppose deeper integration. This insight has also been discussed as the "sleeping giant" thesis (van der Eijk and Franklin, 2004). While voters have increasingly meaningful and consequential preferences about the degree of EU integration, the relevance of these preferences lays dormant because the structure of domestic inter-party competition does not allow voters to express their preferences on EU issues at the ballot box. This gives room for political issue entrepreneurs to challenge the core parties by inserting EU issues into political competition (van der Eijk and Franklin, 2004, 48).[5] The period since 1991 has been described by Marks and Hooghe (2009, 5) as one of "constraining dissensus", with political elites increasingly in need to look "over their shoulders when negotiating European issues".

There is a vast literature on public support for, and opposition to European integration (for a review see de Vries and Hobolt, 2016). Even though there are disagreements over conceptualization and measurement, there is a lot of empirical evidence that suggests that the debate over the degree of European integration increasingly structures the political thinking of Europeans. It seems that European voters have developed relatively stable and well-defined preferences over the degree to European integration. E.g., Boomgaarden et al. (2011) find that mass attitudes towards the EU and European integration are structured by five attitudinal components: affective attitudes, European identity, evaluation of EU performance, utilitarian attitudes and attitudes towards the strengthening of EU integration. The correlation between the factors is fairly high, which might indicate that these factors are sub-dimensions of a latent ideological EU issue dimension. van Spanje and de Vreese (2011) have largely replicated these attitudinal components and find that opposition to deeper EU integration and negative evaluations of the EU's utility have a particular effect on anti-EU voting behavior in EP elections. Maier et al. (2015) come to similar conclusions using an implicit measure of EU attitudes. In a confirmatory factor analysis of mass attitudes towards specific EU policies, Gabel and Anderson (2004) find that citizens preferences are structured by both a left-

---

and try to integrate national sovereignty aspects as much as possible in terms of the left-right dimension. However, traditional parties are internally split on national sovereignty and the cross-cutting nature of territorial and functional coalitions means that traditional parties must fail to fully integrate cross-cutting European issues into the left-right dimension (Bakker et al., 2012, 224).

[5]Building on the literature on issue evolution (Carmines and Stimson, 1989), a small but rich literature has developed that describes and analyzes the process by which political entrepreneurs insert European integration issues into domestic party competition (Hobolt, 2015; van de Wardt et al., 2014; de Vries and Hobolt, 2012).

right and a supranational component. Studies that have investigated the role of European issues at the level of political elites have come to similar conclusions. In a study on expert survey data Bakker et al. (2012) find an EU integration dimension to structure the political space in European countries. Proksch and Lo (2012) offer a word of caution as they find estimates of party positions on European integration as only weakly continuous (see also Marks et al., 2012, for a response). Analyzing EP election party manifestos, Spoon (2012) finds that parties strategically communicate their stances on European integration (see also Braun et al., 2016).

European issues and EP election turnout   Against this broader background, a strand of literature has developed that advocates for European issues as factors explaining turnout in EP elections. This is a challenge to the second-order interpretation of EP elections, since the transfer hypothesis suggests that issues that are specific to the arena play an increased role. The literature can be separated into two lines of research (see Schäfer and Debus, 2015, for this helpful typology). The first line of research remains largely agnostic to the theoretical mechanism that connects EU issues to voting behavior. Rather implicitly it is argued that preferences and attitudes towards the EU capture the degree to which the EU institutions and the European project is seen as legitimate or worthy of support. Analogous to the Voice-Exit literature (Hirschman, 1970), EP election voting behavior then is a way to show or withhold support for the political system. Empirically, the research program relies heavily on the analysis of correlations between EU related factors and aggregate turnout rates. EP election turnout is shown to be higher in countries or regions where support for the EU is greater or which are net receivers of EU subsidies (Stockemer, 2012; Fauvelle-Aymar and Stegmaier, 2008; Flickinger and Studlar, 2007; Mattila, 2003; Studlar et al., 2003; Blondel et al., 1997). The reliance of many studies on aggregate data that leaves them prone to ecological fallacy is certainly problematic. Much needed evidence from individual-level data is still scarce. Based on open-ended interviews, Blondel et al. (1998) showed that a substantial amount of voters at least rationalize abstention in EP elections by negative evaluations of the EU and the EP. Stockemer (2012) corroborates his aggregate level findings with 2009 European Election Study individual-level data and finds a positive effect of EU membership approval on the likelihood to participate in the EP elections. Boomgaarden et al. (2016) report that EU support increased turnout at the 2014 EP election in Austria. These positive empirical findings are however not unchallenged, as evidence to the contrary is also on the record (SchmittEijk2007; Schmitt, 2005; van der Eijk and Franklin, 1996; Schmitt and Mannheimer, 1991). While the majority of the studies suggest that EU related factors might play a role in explaining EP election voting behavior, the findings seem far from robust. Moreover, both empirical analyses based on aggregate- and individual-level data run into severe methodological problems, which will be discussed at a later point in this dissertation.

The second line of research has come closer to developing a coherent theoretical framework to study the connection between EU preferences and electoral participation. A few studies have suggested that abstention in EP elections might be explained by the policy distance between voters and parties on EU integration issue (Hobolt et al., 2009; Hobolt and Spoon, 2012). Building on a multidimensional model of spatial competition, these studies have argued that abstention in EP election is not due to the preferences that voters hold on European issues per se, but rather due to not being represented on these issues by the parties they usually vote for. Particularly, the studies find that voters who voted for a government party in the FOE are more likely to abstain in the EP election if they hold anti-EU integration positions, which the commonly pro-EU integration government parties do not represent. Although the authors fail to make the connection, the notion of lacking representation on policy grounds as a cause for abstention ties in with the literature on policy alienation (Hinich and Ordeshook, 1969; Enelow and Hinich, 1984b). This policy alienation argument was further developed by Schäfer and Debus (2015), however they find no convincing evidence that EU integration preferences play a notable role for abstention via the route of policy alienation.

European issues and vote switching in EP elections Apart from electoral participation in EP elections, the switching of votes from one party in the FOE to another party in the EP election has always been at the heart of the second-order argument. Yet, in a number of studies it has been argued that voter preferences on European issues provide an alternative explanation why voters choose different parties in EP elections than in FOEs. Of course, this is also at odds with the interpretation of the second-order transfer hypothesis.

Various aggregate-level studies have tried to study vote switching by comparing the differences in aggregated election results between FOE and the following EP election. The general verdict of aggregate studies seems to be that Europe-related factors play, if at all, only a minor role in explaining the parties' losses and gains across FOE and EP elections. van der Eijk and Franklin (1991) find that parties primarily acquire votes on the basis of national political concerns also in EP elections. Ferrara and Weishaupt (2004) hypothesize that parties whose platforms give greater salience to European issues and parties that are strongly opposed to European integration do better in European elections than in national elections. In their study of aggregate election results, they do not find empirical support for either, but find that parties that are internally split on European issues do worse in EP elections. In a more extensive study, Hix and Marsh (2007) find that parties that hold Euroskeptic positions did slightly better than average, but only up to the 2004 election. In a follow-up study they find that these parties did relatively worse in the the 2009 EP election (Hix and Marsh, 2011).

As has been widely acknowledged throughout this literature, analyses of aggregate-level election results do not tell much about the motivations of individual voters and run the risk of ecological fallacy. Therefore the literature has converged on

analyses of individual-level survey data. Marsh (2003) and Marsh (2007) showed that voters who think that European integration moves too fast are more likely to desert government parties in EP elections. However, they also find that the party that was switched to in EP elections does not necessarily better represent the switcher's preference on EU issues. Carrubba and Timpone (2005) have argued that voters have different policy preferences at different levels of government. They suggest that individuals who perceive the EP as more powerful are less likely to treat the elections differently. Using Eurobarometer data, they find that voters who are most concerned about environmental issues and perceive the EP as a powerful institution are more likely to vote for a Green party in EP elections. As environmental issues are among the policy issues primarily dealt with at the EU level, they count this as "strong support that individuals actively want different preferences represented at different levels of government" (Carrubba and Timpone, 2005, 279). Analyzing European Election Study survey data, Schmitt (2005) found European factors to play a negligible role in explaining vote switching in EP elections. Sanz (2008) and van Aelst and Lefevere (2011) study concurrent national and EP elections in Spain and Belgium and find that ticket-splitting is at least partially motivated by factors that are specific to the European electoral arena. Clark and Rohrschneider (2009) suggests that electoral choice in EU election is influenced to a considerable extent by how voters evaluate the performance of EU institutions. They find vote switching between FOE and EP elections to be explained by how voters evaluate the EU on its own performance terms, rather than by how voters evaluate the performance of national government.

In a number of studies the role of European issues based on a multi-dimensional model of spatial competition has been investigated. Hobolt et al. (2009) analyzed vote switching of FOE government supporters to different parties in the 1999 and 2004 EP elections as a function of policy distance between voters and the previously chosen party on the Left-Right dimension and EU integration dimension. They find that increasing policy distance on the EU integration increases the likelihood of vote switching. In a similar analysis on the 2009 EP election, Hobolt and Spoon (2012) replicate this finding, and also suggest that the effect is moderated by campaign tone. Hobolt and Wittrock (2011) provide experimental evidence that policy distances on EU integration play a notable role for EP election vote choice, that is however subordinate to policy distances on left-right issues. Moreover, the role of EU integration varies with political information about Europe: When participants are given more information about the party positions on the EU dimension, they put larger the emphasis on this dimension in their decision rationale. de Vries et al. (2011) employs a similar spatial voting model to study EU issue voting in the 2009 EP election. They find that policy distance on EU integration issues predicts EP election vote choice in most EU member countries, and the effect is particularly strong among the politically sophisticated. de Vries and Tillman (2011) and de Vries (2010) use a similar research design to study differences in EU issue voting between Eastern and Western member states, and

the incidence of EU issue voting in FOE respectively. All of the above cited contributions have used cross-sectional data, mainly from the European Election Study (EES). In the more recent past, two studies that use panel data have become available. Giebler and Wagner (2015) compare the vote choice functions of German voters in the 2009 EP and Federal election, and find no evidence that EU integration issues play a different role in the EP election than in the Federal election. Boomgaarden et al. (2016) report that switching of votes from Anti-EU to Pro-EU and vice versa is partially explained by how voters assess EU integration.

2.3 Towards a general model of multi-level spatial voting

In the above, we have seen that there is a growing body of research that suggests that EP elections are somehow (also) about "Europe". However, it is a collection of patchy evidence rather than a concerted research program. So far the literature has not unified behind a common theoretical framework. A substantial part of the literature is not theory-driven and rather advances ad-hoc hypotheses, in which the supposed causal relationships remain opaque. Insofar as theoretical arguments are developed, they are often too specific and fail to be integrated into the wider literature on voting behavior (see also Marsh and Mikhaylov, 2010, 18). As has been observed by Marsh and Mikhaylov (2010, 13), the most promising option to integrate European issues into a coherent theoretical framework of voting behavior is the spatial model of voting. The (Neo-)Downsian framework, and especially the multidimensional model of spatial competition (Davis et al., 1970; Enelow and Hinich, 1984b; Hinich and Munger, 1997), provide ample opportunities to model changing issue dimensions or changing issue saliency between elections.

Substantial advances have been made in this direction: The line of literature arguing from a EU issue voting perspective (Hobolt et al., 2009; de Vries, 2010; de Vries et al., 2011; Hobolt and Wittrock, 2011; Hobolt and Spoon, 2012) has developed a spatial model for the European policy space that allows to analyze the role of European issues for voting behavior in EP elections, and differential voting behavior across FOE and EP elections. The model of spatial competition builds on the idea that voters choose parties with the intent of maximizing utility derived from electoral outcomes. Voters maximize utility by choosing the party whose position minimizes the spatial distance to their ideal point, i.e., they choose the closest match on the relevant ideological preference dimensions. The EU issue voting model starts from the proposition that the relevant issues for European voters can be described in the European policy space. The European policy space consists of a Left-Right dimension, that captures policy considerations related to the traditional line of political contestation, and a EU integration dimension, that captures considerations related to the conflict over more national independence versus more EU integration.

The EU issue voting model is an excellent basis for the integration of the Europe Matters argument into a coherent framework, however I find that this potential has not been fully realized. There are important implicit assumptions, theoretical gaps and fallacious simplifications in previous contributions that I will point out in the following. By addressing these concerns, a complete model of spatial voting over multiple electoral arenas and dimensions will arise that allows for a clear formulation of the theoretical points of contention and how these can be empirically tested. Instead of relying on a mathematical exposition, I shall try to formulate the theoretical points of contention verbally and graphically.

In a two-dimensional policy space such as the European, a key factor that determines spatial utility is the saliency or weight that voters assign to each of the dimensions (Davis et al., 1970; Feld et al., 2014). Applications of the EU issue voting model have so far used estimates of the weight parameters somewhat pragmatically to draw inferences about voting behavior. While it is correct that if voters assign a positive weight to policy distances on the European integration dimension in an election (indicated by a significantly negative coefficient), one can conclude that voting behavior is influenced by European considerations in this election, other more far-reaching conclusion do not automatically follow. As I will show in the following, it can not be inferred from a significant coefficient that the transfer hypothesis does not hold, or that differential abstention or vote switching between the FOE and EP election is motivated by EU integration considerations.

The recalibration hypothesis

In order to pose a challenge to the second-order model's transfer hypothesis, the research question that the Europe matters argument constitutes, needs to be clarified first. First of all, the transfer hypothesis does not preclude a role of European issues in EP elections. It only suggests that issues remain the same across FOEs and SOEs. So if European issues drive voting behavior in EP elections, this would be fully consistent with the second-order model - if European issues play an equal role in FOE as well. Therefore, the "more pertinent question is whether European issues play a greater role in European Parliament elections than in national elections"(Marsh and Mikhaylov, 2010, 13).

Accordingly, it is not sufficient to show that European issues play a role in EP elections (or for vote switching in EP elections). Arguing from the spatial model, voter ideal points and party positions should be thought of as independent of the electoral arena. This is also how the EU issue voting literature has conceptualized preferences.[6] If voter ideal points and party positions are fixed across elections, changes in voting behavior across elections can only be caused by changes in the weight that voters assign to the two dimensions (Feld et al., 2014). So to explain

---

[6]An exception is Carrubba and Timpone (2005), where voters have different preferences at different levels of government. However, their argument is built on a balancing model, not a spatial model.
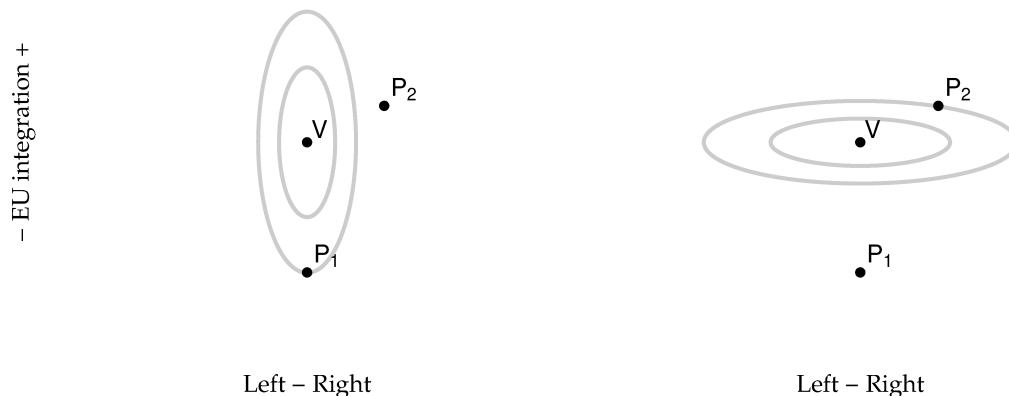
Figure 2.1: Illustration: First Preference for a different party in FOE (left) and EP election (right) due to changing dimensional weights

a change in voting behavior between elections that is motivated by European spatial considerations, the weight assigned to distances on the European integration dimension needs to be higher in EP elections than in national elections.

The spatial voting scenario in Figure 2.1 illustrates how changing weights can motivate vote switching. In the FOE (left panel), the voter prefers party P_1 over party P_2 because P_1 is located at a higher utility indifference contour line. In the EP election (right panel), the voter prefers party P_2 over party P_1, as P_2 now has a lower *weighted* spatial distance to the voter's ideal point. This is not because a change in the voter's ideal point or changes in the parties' positions, but only due to the increased relative importance the voter assigns to policy distances on the EU dimension. So it is not that the voter that moves, but rather the shape of the policy space that changes across electoral arenas.

That voters care more about specific issue dimensions at different levels of government is *the* lynchpin of the EU issue voting explanation. But is this empirically the case, do voters "recalibrate their issue space" according to the electoral context (Marsh and Mikhaylov, 2010, 13)? Unfortunately, this has not been shown so far, although some authors have recognized the importance of this question (e.g., Hobolt et al., 2009, 97). What has been shown is that policy distance on EU integration explains voting behavior or vote switching in EP elections. However, in order to establish the validity of the recalibration hypothesis empirically, this is only sufficient if and only if one is willing to assume a priori that European issues play no role in FOE. Such an assumption would however be at odds with a substantial literature, partially by the same authors, that has shown that European issues indeed play an increasing role also in national elections (Evans, 1998; Evans, 2002; Gabel, 2000; Tillman, 2004; de Vries, 2007; de Vries et al., 2011; Schoen2008; Hobolt and Wittrock, 2011). The strong institutional interlocking of domestic and

European polity and policy indeed provides ample reasons to think that the process of Europeanization has also permeated domestic political competition. As "government officials elected through national elections participate in the EU Council of Ministers [and] elected government leaders directly represent the interests of their member states and their citizens in the European Council" (de Vries, 2010, 92), policy-motivated voters that hold relevant opinions on European issues should take these into account when they vote in national elections.

In Chapter 5, I will show how the recalibration hypothesis can be precisely formulated and tested in the multidimensional spatial voting framework. Voters can be said to put a higher emphasis on EU issues in EP elections if the relative dimensional weight of European issues is larger than in national elections. I find evidence that this is likely to be the case in the data under investigation. By transforming the recalibration hypothesis from an implicit a priori assumption into an empirically testable hypothesis, the EU issue voting framework is put on on a solid theoretical and empirical footing to explain differential voting patterns across FOE and EP elections.

EU issues and differential turnout

The EU issue voting argument also entails a clear implication about differential voter turnout in EP elections that has not been fully recognized.[7] The spatial theory of voting identifies two policy-related reasons for voters to abstain from voting: Alienation and indifference (Hinich and Ordeshook, 1969; Enelow and Hinich, 1984b). "Voters abstain from alienation if their utility for their favorite candidate fails to exceed what they consider a minimum level. [. . .] A voter abstains from indifference if the utility difference between the two candidates fails to exceed some minimal amount." (Enelow and Hinich, 1984a, 464). In other words, voters are more likely to abstain if they do not feel represented by any of the available parties, or if they are located at equal policy distance to two parties.

Some contributions have implicitly argued that abstention in EP elections may arise from voters not being represented by their traditional parties on European issues (Hobolt et al., 2009; Hobolt and Spoon, 2012).[8] More precisely, government supporters that disagree with the usually pro-European integration positions their parties take, are more likely to abstain in EP elections, after having turned out in the FOE. So the argument is about the switch from participation to non-participation, to which I will refer as differential abstention in the latter. How can alienation explain differential abstention between elections? First of all, if voter ideal points, party positions and the dimensional weights remain fixed across the

---

[7]An exception is Schäfer and Debus (2015), however they are not interested in explaining differential turnout, but EP election turnout.

[8]There also might be an argument based on indifference, as centrist parties take on similar positions on European integration issues, but this argument seems less relevant and shall not be pursued in the following.

elections, there is no policy-related reason to expect differences in turnout across elections. Voters would be just as alienated from parties in FOE as in EP elections. Again, an explanation for differential abstention based on the spatial model necessarily requires that voters recalibrate: The relative weight of the policy dimensions needs to change across elections in order to motivate different voting behavior across election.

If EU integration issues indeed play a larger role in the spatial calculus of voters in EP elections than in national elections, i.e., the recalibration hypothesis is fulfilled, a mismatch between party positions and voter preferences on EU integration issues can lead to a representational gap that is larger in EP elections than in FOE. This gap can then motivate more abstention due to increased policy alienation in the recalibrated space. The "policy alienation due to recalibration" argument implies that which voters feel more alienated fundamentally depends on the distribution of parties in the European policy space. Voters can only feel alienated if the party system does not offer political alternatives that represent the recalibrated voter policy preferences. Ultimately this means that the relation between European issue disagreement and differential turnout might vary a great deal between different party systems and electoral contexts. While empirical research clearly needs to carefully account for the features of party systems, this does not absolve me from proposing a general theoretical model.

The most widely accepted model of party-competition in the two-dimensional European policy space among European election scholars is the inverted U curve (Hooghe et al., 2002; Rohrschneider and Whitefield, 2016). It states that left-right centrist catch-all parties generally adopt similarly pro-European positions, while left-wing and right-wing parties, even if for different reasons, take more anti-EU positions. Given such a constellation of party positions in the European policy space, the voters with high policy alienation in EP elections can be identified. Firstly, these are left-right centrist voters with anti-EU preferences. As centrist parties are also pro-European, these voters would have to sacrifice too much policy closeness on the left-right dimension to find a matching anti-EU party. Secondly, pro-European voters that hold extreme positions on the left-right dimension. These voters do not feel represented by the anti-EU left- and right-wing parties on European issues and would have to give up too much on left-right closeness in order to get to their preferred position on European issues, which are all located in the center.

To illustrate how changes in dimensional weights - EU integration issues becoming more important relative to Left-Right issues in the EP election than in the FOE - can lead to increased policy alienation in EP elections, I simulate party utilities as a function of EU integration and Left-Right distances for possible voter ideal points in the European policy space. Policy alienation then is proportional to the minimum party utility (among the available parties). I generate four parties that are aligned along an inverted U-curve: an anti-EU Left party, a pro-EU center-left Social Democratic Party, a pro-EU center-right Conservative Party, and an
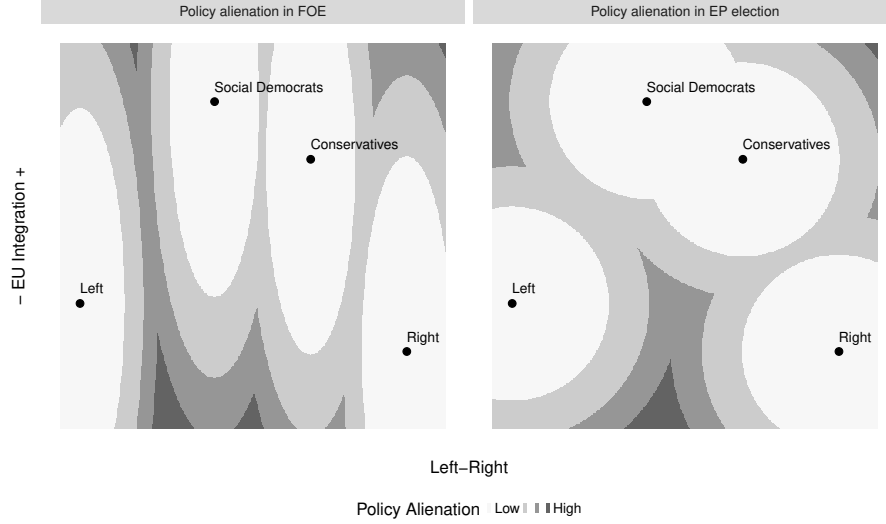
Figure 2.2: Policy alienation heat map: Weight of EU integration distances 5% in FOE, 50% EU in SOE

anti-EU Right party.[9] In the FOE, EU integration distances are assumed to play a subordinate role to Left-Right distances (weighted 5% and Left-Right distances 95%), in the SOE both dimensions are weighted equally (50%/50%). Figure 2.2 plots the policy alienation that a voter located at a particular point in the European policy space is exposed to, in the FOE scenario (left panel) and the SOE scenario (right panel). Darker colors indicate more policy alienation.

Since party utility is mainly about Left-Right distance in the FOE, most voters feel well represented on these issues because the available parties are evenly distributed along the Left-Right dimension. Lacking representation on EU integration is largely inconsequential due to the low weight that voters assign to the dimension in FOE. As the weight of the EU integration dimension increases in the EP election, representational gaps occur since the particular constellation of the parties in the two-dimensional space do not represent certain voter groups. These

---

[9]Spatial utility of voting for party j in the FOE is calculated by $U_{ij} = -.1(p_{j_{eu}} - v_{i_{eu}})^2 - 1.9(p_{j_{lr}} - v_{i_{lr}})^2$, and utilities in the SOE by $U_{ij} = -1(p_{j_{eu}} - v_{i_{eu}})^2 - 1(p_{j_{lr}} - v_{i_{lr}})^2$.

Figure 2.3: Policy alienation heat map: Difference between FOE and SOE alienation due to change in weight from 5% EU to 50% EU.

will be less likely to turn out in the EP election due to policy alienation. However, in order to explain differential abstention, i.e. voting in the FOE and not voting in the EP election, the difference in policy alienation between the elections is decisive. This is plotted in Figure 2.3. The voter groups most affected by increasing policy alienation due to the change in weights are voters with Left-Right centrist anti-EU integration preferences and left- and right-wing pro-EU voters. In this scenario, one should therefore find that being more in favor of EU integration should have a negative effect on differential turnout among left-wing voters, a positive effect among centrist voters, and again a negative effect among right-wing voters.

To summarize: A spatial explanation for differential abstention due to policy alienation requires that voters recalibrate their issue space across elections. If voters recalibrate, not all voters are equally affected by an increase in weight of the EU integration dimension in EP elections. Policy-related alienation increases conditionally on the left-right preferences of a voter, and the distribution of parties in the European policy space. The inverted U-shape model of party competition posits that there may be a typical representational gap in many member states. Following the left-right continuum, pro-EU leftist voters, anti-EU centrist voters and pro-EU rightist voters should be more likely to abstain in EP election due to increased policy alienation.

In Chapter 4, I empirically test the argument that differential abstention in EP elections can be explained by increased policy alienation as a consequence of a recalibration of the issue space. Rather than modeling the differences in policy alienation directly, I choose an indirect route. As I have shown here, each party system will generate specific representational gaps if dimensional weights

27

are recalibrated. The increased policy alienation that these representational gaps create imply a specific pattern, or functional form, in which specific combinations of EU integration and Left-Right will lead to differential abstention. From the party system under investigation I draw observable implications about the trivariate relationship between EU integration and Left-Right voter ideal points and test to which this supposed relationship is consistent with the data.

Policy-motivated vote switching

As noted above, if voters recalibrate their issue space between elections, this might lead to different vote choices in the FOE and EP election. However, this is inherently contingent upon the "supply side", the policy positions that alternative parties offer to the voter who is at risk of switching away from her previous party. Only if there is a party that better matches the reweighted multidimensional policy preferences in the EP election will a voter switch to a different party. This vital part of the switching process, that there is not only a party from which a voter switches away, but also a party a voter switches towards, has been woefully omitted in previous studies that try to explain vote switching as a function of EU integration preferences (e.g. Clark and Rohrschneider, 2009; Hobolt et al., 2009; Hobolt and Spoon, 2012; Hobolt, 2014; Boomgaarden et al., 2016).

The problem is not only of theoretical nature, but casts doubt on a popular kind of research design that is commonly employed to study differential voting behavior across FOE and SOE. I find that, although the EU issue vote switching literature has developed its theory based on a spatial model, their research hypotheses do not necessarily follow from the spatial model. How can this be? First of all, the spatial theory of voting is a theory of choice among different alternatives, i.e. vote choice is the dependent variable. Most EU issue voting contributions however seek to explain vote switching, that is the change of party vote choice between FOE and EP elections. The hypothesis to be tested in vote switching studies is often formulated as follows. "The greater the distance on the issue of European integration between voters and the party they previously supported, the greater the chance of defection [. . .] at EP elections." (Hobolt et al., 2009, 97). Unfortunately, this hypothesis is only a valid deduction from spatial theory if specific circumstances are met. This is because in the spatial model, the probability to vote for a different party is inherently conditional on the positions all available party alternatives take. Simply put, if the party a voter has voted for in the FOE is the most anti/pro-European party there is, being more anti/pro-European than this party will not increase the chances of defection in the EP election. In fact, spatial theory would suggest the opposite - that it will make a vote for that same party more likely.

Figure 2.4 illustrates that modeling the likelihood of vote switching as increasing with increasing ideological distance to the previously supported party is inconsistent with the theoretical predictions of a complete spatial model. I calculated
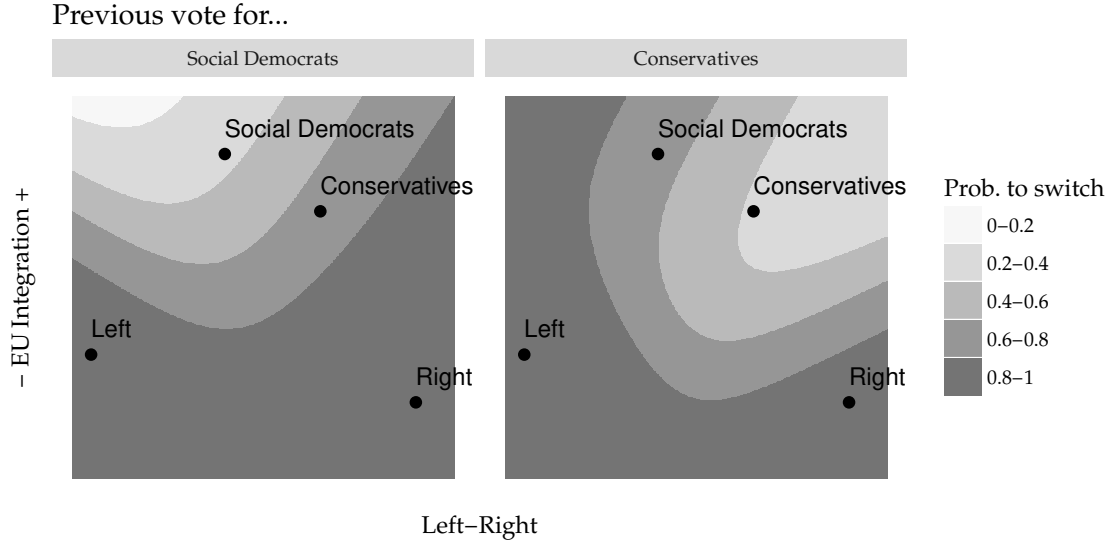
Previous vote for...

Figure 2.4: Illustration: Heat map of spatial vote switching probabilities given previous vote for Social Democrats (left) or Conservatives (right)

the switching probability for all possible voter ideal points in the policy space according to a standard spatial voting model. The scenario entails four parties in the constellation encountered in the previous section.[10] One can clearly see that the probability of vote switching does not symmetrically increase with increasing distance from the previously chosen party. It is even decreasing with distance in a direction where there is no competition from other parties. E.g., a former Social Democratic voter (left panel) is predicted to be more loyal to his party if he is more pro-EU and leftist than his party. This is because there is no alternative party in this direction that would better match his preferences. The Conservative Party (right panel) is able to hold voters that are centrist on both dimensions because there is no other party that is located closer to the multidimensional center. However, the Conservative party is more vulnerable to competition from the the Populist Right on EU integration issues.

As this illustration shows, it is not true that the probability to switch increases monotonically with distance to the previously chosen party's position. Vote switching probabilities are inherently dependent on the constellation of parties in the policy space (and the relative weight voters assign to the dimensions). Relatively complex spatial patterns can evolve from simple models of party competition. To

---

[10]Spatial utility of voting for party j is $U_{ij} = -1(p_{j_{eu}} - v_{i_{eu}})^2 - 1(p_{j_{lr}} - v_{i_{lr}})^2$ i.e., voters assign equal weights to left-right and EU integration distances. Party choice probabilities are calculated from a conditional logit model (McFadden, 1974). The probability to switch then is 1 minus the probability of voting for the same party again.

be fair, this is certainly known to all the authors and was partially acknowledged in the original contributions. E.g., Hobolt et al. (2009, 98) point out that "defection on the basis of issue voting requires that another party offers a position closer to the voter's ideal point". Given this "confession", the authors limited their analysis to government party voters, whose parties are usually the most pro-European, and additionally reformulate their hypothesis as directional, in the sense that "Voters are more likely to defect [...] if they are less supportive than their party of European integration. While these simplifications might have saved the theoretical integrity of the original hypothesis, they inevitably reduce the scope of the theory to a specific partisan subset of the electorate. This limitation means at the same time that this modified theory is ill-suited as a general model of EU issue voting.

Taking a step back, the core theoretical problem of many studies of EP vote switching is the mismatch between theory, hypotheses and models. As was laid out above, this becomes clearly visible in the invalid blending of vote choice and vote switching.[11] Rather than addressing this fundamental problem, the literature that seeks to explain changes in voting behavior across FOE and SOE elections has converged on the development of more elaborate, or on a more critical note, convoluted classification schemes for voting patterns in recent years (e.g., Weber, 2011; Hobolt and de Vries, 2016). While classifications can be theory-driven, they tend to invite arbitrary classification choices and ad-hoc hypothesizing about specific groups. On a more general note, classifying voting behavior across multiple elections and using the classifications as the dependent variable stands in stark contrast to the methodological approaches pursued in other fields of political science research that describe and model change over time. Political science has since its early days understood change in nominal variables as Markov processes (e.g., Converse, 1964; Dobson and Meeter, 1974), where the state at one point in time is conditional on the preceding state, and the probability of change in states over time is expressed in terms of transition probabilities. Although the second-order model and the Europe Matters argument is intrinsically concerned with changes in voting behavior between FOE and SOE, there exists no contribution that has conceptualized this as a Markov process. This means that the methodology of the literature on EP elections is disconnected from the academic mainstream, and fails to profit from the vast literature on dynamic discrete choice and Markov models. To point this out very clearly: This is not primarily a statistical issue (although there are very good statistical reasons not to classify). The problem I see is that theorizing on the basis of classifications is fallacious. Individual voters choose which party to vote for among a set of parties, once in the FOE and once in the

---

[11]A number of exceptions to this patterns can be identified, but these are mostly confined to studies of voting in a single policy arena. Hobolt and Wittrock (2011) analyze multi-party vote choices in an experimental situation as the dependent variable, de Vries (2010) and de Vries et al. (2011) analyze multi-party vote choice in national, respectively EP elections. As an exception to the exception, Carrubba and Timpone (2005) analyze the vote for the Green or government party as binary choices across FOE and EP elections.

EP election.[12] 'Switching' is not on the ballot paper. It is only the researcher who classifies the voting sequence as 'switching'. Therefore theories that work on the level of classifications (e.g., vote switching vs. partisan voting), will inevitably come into conflict with any general theory of the individual voter. While there might have been reasons to classify because of data limitations in the past, I find that this is not the way forwards. In my opinion, the classification approach is a methodological dead end that should be abandoned.

In Chapter 7, I develop an alternative approach that allows for the study of vote switching based on a coherent framework of spatial voting. Instead of relying on classification schemes, I model changes in party vote choice across elections as a Markov process. The effect of distances on EU integration issues on transition probabilities is then studied in a dynamic discrete choice framework.

---

[12]The EP election vote can of course be modeled as conditional on the FOE vote. This is what second-order theory is about.

Chapter 3

Methodological challenges and the promises of panel data

The debate between second-order theorists and Europe matters proponents, which has structured the entire field of research over the last decades centers on the question whether electoral outcome, or differences between electoral outcomes in FOE and EP elections, are explained by EU related factors. On both sides of the argument, the goal of empirical research has so far been to establish either the absence or presence of correlational patterns between electorals outcome and EU related factors in observational aggregate-level and cross-sectional individual-level data. A remarkable feature of this important debate is that, unlike in other fields of research, it has not been a debate over methodology. Proponents of both sides of the argument have used similar research designs, however tend to come up with different answers. The discrepancy in results then tends to be explained using largely ad-hoc arguments about the reasons why a coefficient is significant for this country/election/year/party, or by proposing additional moderating factors. Consequently, a cottage industry of academic papers which ask each electoral cycle whether a particular EP election is "still second-order" or "finally about Europe" has developed. In my opinion, this is not the way forward. If anything, the state of methodology in the field of EP election research needs critical reflection. I find that both strands of the literature have displayed little care for methodological issues that are prone to undermine the validity of the drawn inferences. Moreover, there is remarkably little cross-pollination of methodology from other fields of research. Some of the methodological problems are certainly grounded in paramount data limitations. However, better data, multi-election panel data in particular, is beginning to become available. Unfortunately, the state of methodology in EP election research indicates that the field is unlikely to be going to make the best of this new data.

Rather than subjecting the relevant contributions to a systematic methodological strip-search, what I want to offer in this Chapter is a fresh start. Second-order theorists and Europe matters proponents alike have remained largely oblivious,

or at least silent in their contributions, about causal identification. So I think there is some benefit in clarifying the fundamental inferential problem first that arises in the analysis of voting behavior between FOE and EP elections, or for that matter, multi-level voting behavior in general. In a second step, I point out how panel data can help not only by providing better measures, but also by allowing to control for unobserved heterogeneity, the most likely danger to the validity of inferences in voting behavior research. Eliminating potential heterogeneity bias however requires the employment of more powerful statistical techniques such as fixed effect models. These come with their own set of challenges for discrete voting behavior data. An attractive option is the conditional likelihood approach to panel data, which not only the field of voting behavior research, but political science research in general has so far woefully overlooked. I discuss conditional fixed effect logit models for choice data, and how they can be employed to exploit the advantages of political panel data to the fullest. Following, I sketch out the research designs that I will employ in Chapters 4-7 and discuss their properties. Finally, I discuss case selection and introduce the Making Electoral Democracy Work (MEDW) Bavaria panel data set and explain the measures that I will use in the empirical part of this dissertation.

3.1 The fundamental problem of causal inference in multi-level voting research

All research that seeks to compare voting behavior across elections at different levels of government faces a common inferential problem. Changes in the level of government at which the elections take place are in most instances accompanied by changes in time. Simply put, EP elections usually take place at a different point in time than FOE. This means that it is practically impossible to credibly distinguish whether differences in voting behavior between elections are caused by the change in the electoral arena, or by the change in time. Clearly, this problem applies to data on individual as well as aggregated voting behavior. Using the language of causal inference, the treatment that is applied to voters is that the electoral arena changes. Defining the FOE as the control group, and the EP election as the treatment, all voters are either in the control group or in the treatment group at one point in time.

Figure 3.1 illustrates the problem. In a typical situation, voting behavior is only observed in the FOE at time point $t = 1$ and the EP election at $t = 2$. The potential outcomes (Rubin, 1978) of an EP election at $t = 1$ and of a FOE at $t = 2$ (Y_{EP_1} and Y_{FOE_2}) are not observed. This problem is unlikely to be solved in the field of multi-level voting behavior research. A substantial proportion of the causal inference toolbox is simply not available. Controlled randomization of real-world elections is obviously impossible. Laboratory experiments can provide valuable insights into EP election voting (Hobolt and Wittrock, 2011; Blais et al., 2016). However,
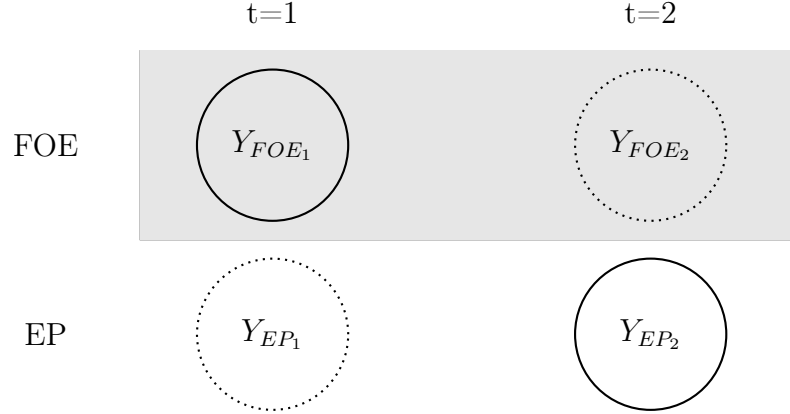
Figure 3.1: Observed (solid line) and Unobserved (dashed) potential outcomes

external validity is often reasonably doubted due to the difficulty to recreate a realistic electoral context in the lab (Lupia, 2002; Lupia and McCubbins, 1998; McDermott, 2002). Natural experiments[1] exist in the form of concurrent elections, cases in which multiple elections at different levels of government are held at the same point in time. While they can be employed to great effect to provide causal answers to important research questions (e.g., Leininger et al., 2016; Fauvelle-Aymar and François, 2015), they are available only in special situations where only a subset of the population is treated and treatment assignment is plausibly as-if random. Natural experiments are by design not available for a comparison of FOE and EP election voting behavior, since both elections are applied at the national level at minimum. Therefore, the entire electorate is either treated or not. Concurrent FOE and EP elections are not a solution. If the FOE and the EP election are held on the same day, all subjects belong to the treatment and the control group at the same time.[2] It can not be made plausible that the FOE and EP election outcomes, generated by a concurrent election, are viable substitutes for the missing potential outcomes if they are regarded in isolation. EP voting behavior given a concurrent FOE election is unlikely to be the same as EP voting behavior in a (counterfactual) single EP election, because votes tend to be sticky across ballots. This has been discussed in the context of cross-ballot and cross-election contamination or interaction effects (Herron and Nishikawa, 2001; Ferrara et al., 2005). Theoretically, contamination in both directions is possible. However, it seems plausible that EP election behavior is contaminated more by the more important FOE than the other way around. The election campaign should also

---

[1]I use the term here in its widest sense (Dunning, 2008, 2012). For recent uses of natural experiments in voting behavior research see e.g., Persson et al. (2013) and Ferwerda (2014). For a critical note on the term and usage see Sekhon and Titiunik (2012).

[2]Expatriates that are only allowed to vote in the EP but not in the FOE are an exception that could potentially be leveraged in a Difference-in-Difference design. However, expatriates clearly do not constitute a randomized or balanced control group, and parallel trend assumptions might be dubious at best.
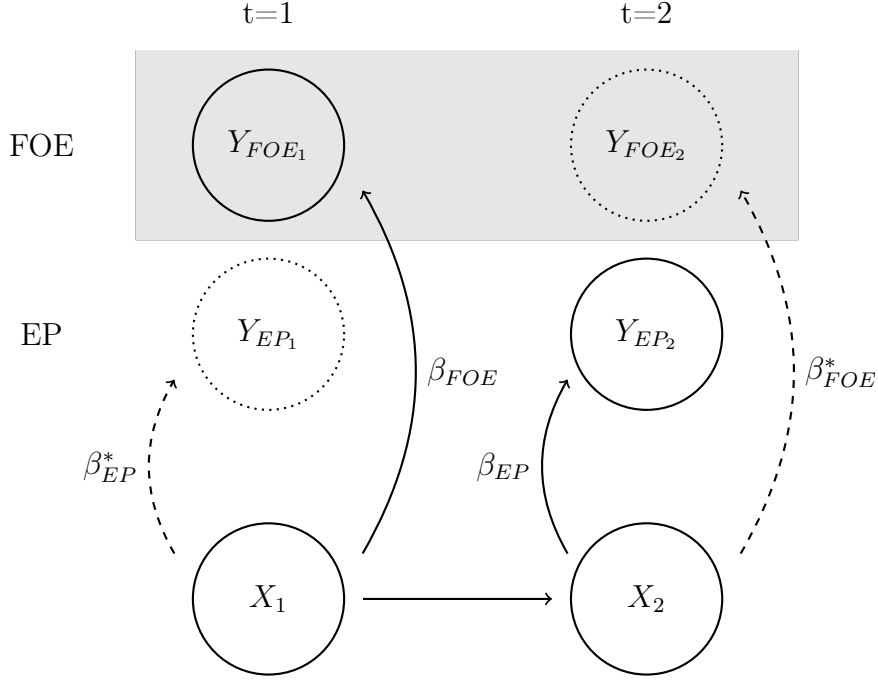
Figure 3.2: Potential effects

be thought to be dominated by domestic politics, crowding out European aspects. An EP election attached to another election should therefore differ substantially from a single EP election. While important in their own right, the effects reported by studies of ticked-spliting in concurrent EP elections (van Aelst and Lefevere, 2011; Sanz, 2008) are likely minimal effects.

Although the above discussion does not explore the causal inference toolbox exhaustively, it should become clear that one is unlikely to find a good-enough replacement for the missing potential outcomes when studying changes in voting behavior in multi-level elections. Thereby we need to give up hopes of "clean and simple" causal identification strategies. Observational data, on which the relevant literature has so far almost exclusively relied on and to which this dissertation makes no exception, offers little leverage on the fundamental problem of causal inference. To even come closer to solving the problems, the data requirements are prohibitive. Intuitively speaking, to even somewhat plausibly disentangle change in electoral arenas and change in time, one would need to observe voting behavior of the same voters across multiple FOE-EP elections sequences. Such data is simply not available so far. In that light, a causal interpretation of estimates can only be saved by making untestable assumptions.

To see what the assumptions are that I need to make to allow for a causal interpretation, I further develop Figure 3.1 to incorporate the independent variable of interest in this dissertation, EU-related policy preferences, in Figure 3.2. Loosely borrowing from the method of directed acyclic graph (DAG) (Pearl, 2009), ob-

served variables are indicated by a solid circle, solid arrows represent a possible existence of a causal relationship. Dashed line indicate unobservables. Clearly, preferences may vary over time, so let x_1 represent policy preferences at the time of the FOE, and x_2 at the time of the EP election.[3] Let β_{FOE} denote the causal effect of policy preferences x_1 of FOE voting behavior, and β_{EP} of x_2 on EP voting behavior. β^*_{EP} denotes the causal effect of policy preferences x_1 on the unobserved potential outcome EP voting behavior at time point 1, and β^*_{FOE} of x_2 on potential FOE voting behavior at time point 2. Clearly, β^*_{EP} and β^*_{FOE} cannot be estimated.

To reiterate, the central research question this dissertation seeks to answer is whether the effect of policy preferences on EU-related issues on EP election voting behavior is different from their effect in FOE elections. Looking at Figure 3.2, this translates into estimating $\beta^*_{EP} - \beta_{FOE}$, or $\beta_{EP} - \beta^*_{FOE}$. However, one can only observe β_{FOE} and β_{EP}. Under which circumstances does $\beta_{EP} - \beta_{FOE}$ capture the true quantity of interest? Clearly, if $\beta^*_{EP} = \beta_{EP}$ and $\beta^*_{FOE} = \beta_{FOE}$, meaning if the potential effects do not vary over time. This seems a strong assumption to make. However, it closely resembles an assumption that is generally implicitly in analyses of panel data or time-series: That the regression coefficients of time-varying variables are constant over time (Harvey, 1978). In this light, the assumption seems less problematic. Furthermore, case selection can help to make this assumption more trustworthy. Ideally, one would want to study a case where the time period between FOE and EP election is small and politically uneventful. This would make it more plausible to believe that the causal relation to the unobserved potential outcome remains stable. However, as suggested in the initial discussion, there should be at least some time between the elections in order to prevent possible contamination.

Based on Figure 3.2, one can also clarify a research design that has been used to study differences in FOE and EP voting behavior so far. While the limitations of cross-sectional data are one issue, as will be discussed below, I think there are more fundamental problems with the research design. The research design can be written up as $\Delta(Y_{EP}, Y^*_{FOE}) \sim f(\beta x_2)$, where Y^* indicates Y as recalled by the respondent at time point 2, and f() is some function, usually the logit. I see two main problems with such a research design. First of all, the research design uses a measure of change directly as the dependent variable (indicated by Δ). E.g., recalled participation in the FOE and non-participation in the EP election is coded as 1 and participation in both elections as 0, or recalled vote choice in the FOE and EP vote choice for a different party is coded as 1 etc. At least to me, it remains completely opaque what this "first-difference approach to discrete

---

[3]I assume that preferences are a characteristic that is intrinsic to the voter, and independent of electoral arena. This means that, a specific point in time, a voter has only one ideal point, and not one FOE ideal point and a different EP election ideal point. While this understanding of preferences is consistent with how preferences are defined in the spatial voting literature, it may conflict with theories of vertical policy balancing (Fiorina, 1996; Kedar, 2006; see also Carrubba and Timpone, 2005, for policy balancing in EP elections).

voting behavior" shall accomplish in terms of identification. None of the many contributions pursuing this approach has provided any reason for doing so. Since there is no first-difference design for limited dependent variables (unlike in the linear model), I conclude that it does not serve any purpose. It seems this is merely done because that is the way it was previously done in the field of research of EP election voting behavior. However, this is generally not done in virtually all other fields of empirical research that are concerned with discrete change. My second issue with this research design concerns β. It is not clear to me what β really is. Sure, β relates x_2 to the probability of the dependent variable being one. But is this what we are interested in? I would suggest not. Again, the research question I (and the Europe matters literature in general) seek to answer is whether EU-related issues play a more pronounced role in EP elections than in other elections. I.e. whether there is a difference in effect sizes between electoral arenas. Estimating the effect of policy preferences on the change in voting behavior is of little help.[4]

3.2 Making the most of panel data

One reason why answering the research question ultimately requires panel data can be seen in Figure 3.2. If policy preferences are understood as time-varying as I believe they should, policy preferences need to be measured ideally at the point in time the different elections take place. Otherwise one does not stand a chance to distinguish the quantity of interest, change in the *effect* of policy preferences and voting behavior across electoral arenas, from temporal change in policy preferences between elections. The vast majority of previous individual-level studies have used cross-sectional survey data, usually from the European Election Survey (EES). The ability of cross-sectional studies to measure the relevant concepts is limited. Studies taking place at the time of the EP election have to infer voting behavior in the preceding FOE elections using so called vote recall items. Vote recalls record FOE voting behavior as remembered by the respondent, often multiple years after the FOE has taken place. Unfortunately, voters are surprisingly bad a remembering previous voting behavior correctly (e.g., Weir, 1975; Waldahl and Aardal, 2000; van der Eijk and Niemöller, 1979). This may have consequences for the integrity of causal inference about the role of preferences for voting behavior. If the likelihood of misreporting is associated with policy preferences, this will confound the β estimates. However, there is not much one can do about this with cross-sectional data. Moreover, past policy preferences can hardly be inferred

---

[4]I think a better design for cross-sectional individual-level data would be to pool the observations and determine the difference in β using an interaction effect. $Y \sim f(\beta_1 x + \beta_2 D \times x)$, where D is a dummy indicator signifying whether an observation stems from the EP or not. β_2 then expresses the difference in the effect of x given D. Alternatively, a bivariate probit setup, where choice in both elections are predicted jointly like in Carrubba and Timpone (2005) is an attractive option, as it allows for errors to be correlated between the elections.

in cross-sectional survey studies. The only information on preferences therefore stems from measures conducted at the point in time the survey took place, usually a couple of days before the EP election. This means that in order to answer the research question, one has to assume that measures of present policy preferences reflect past preferences, i.e., to assume that policy preferences are time-invariant. Although this might be a workable assumption, it is an inconvenient one from the standpoint of causal inference. With cross-sectional data, the alternative explanation that observed changes in voting behavior are the consequence of changing voter characteristics and not the consequence of changing the electoral arena can never be confidently excluded.

Unobserved heterogeneity   That the field of EP election research is in dire need of panel data that includes more than EP elections is widely undisputed. Only one indication is that the flag ship of the research community, the European Election Studies project launched an initiative to extend its 2014 EP election Voter Survey by an online panel component that will survey respondents from eight countries again in the subsequent FOE. This certainly bodes well for the future. But some studies that use panel data have already become available recently. Lefevere and van Aelst (2014) use the Dutch LISS Panel, and Boomgaarden et al. (2016) use the Austrian National Election Study to study changes in voting behavior between the respective FOE and the 2014 EP election, while Giebler and Wagner (2015) employed the German Longitudinal Election Study to study voting in the 2009 EP and 2009 FOE in Germany. However, I find that these studies have not exploited the full potential of panel data as they used the panel data only insofar as it provides more timely measures of voting behavior and voter characteristics.

Panel data offers much more than better measures. A key motivation for using panel data is the ability to control for all time-invariant factors without the need of observing it. Individual-specific time-invariant factors, which may be understood as individuals' permanent innate ability or inclination to produce a certain outcome, are commonly referred to as unobserved heterogeneity.[5] Social science research, belonging to the population sciences, cannot afford to ignore the paramount variation between individuals (Xie, 2013). Therefore, unobserved heterogeneity is generally considered the most relevant factor undermining causal inferences in the social sciences. "The problem of causal inference is fundamentally one of unobservables" (Halaby, 2004, 508). Unobserved heterogeneity threatens the integrity of causal inference if it acts as a confounding factor. A confounding factor is one that is causally related to both the dependent and the independent variable. If the confounder is not controlled for, a biased estimate of the causal effect between independent and dependent variable is obtained. This is generally referred to as heterogeneity, omitted variable or specification bias. Figure 3.3 illustrates

---

[5]The adjective "unobserved" is somewhat misleading, as it conceptually includes, amongst the factors that are not measurable a priori also all factors that remain unspecified or are omitted from an analysis, irrespective whether they are measured or not.
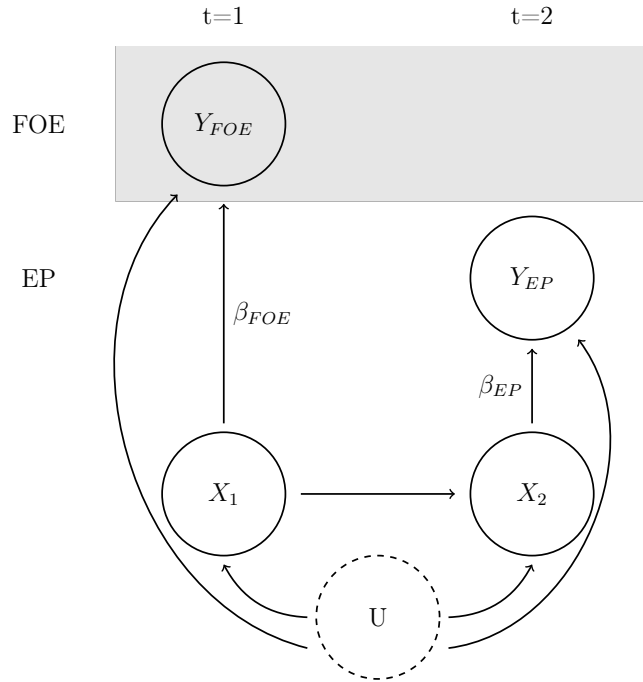
Figure 3.3: Confounding due to Unobserved Heterogeneity

the problem. Unobserved factor U is related to both X and Y in both elections. Failure to control for U will lead to biased estimates of β_{FOE} and β_{EP}.

Three main strategies to deal with heterogeneity bias can be distinguished. The traditional route is to find a parametric approximation for the heterogeneity by including control variables, to the degree that the remaining unexplained variance is purely random error. While most researchers are generally aware of the futility of such an undertaking given the state of social science theories and data (Andreß et al., 2013, 242), in the absence of panel data it is often the only way to potentially reduce bias. Even if researchers identify relevant control variables and specify the functional form correctly, the possibility that there exist yet unobserved confounding factors can never be excluded. By design, only factors that are observed can be controlled for. From a viewpoint of causal inference, the "statistical control" strategy is least likely to produce causal estimates. If panel data is available, two superior strategies become available: Random effect and fixed effect models. They are superior to the statistical control strategy, and related techniques such as propensity score matching, because they do not only control for observed, but unobserved heterogeneity as well. The defining difference between random and fixed effect models are their assumptions about the relation between unobserved heterogeneity and observed independent variables (Mundlak, 1978; see also Wooldridge, 2002, 254). Random effect models model heterogeneity as a random variable that is assumed to be unrelated to the explanatory variables. Fixed effect models, on the contrary, allow observed explanatory variables to be related in any way to un-

observed heterogeneity. Since unobserved time-invariant factors are allowed to be correlated in any way with the independent variables, fixed effect models allow to control for all potential confounding due to unobserved heterogeneity. This feature makes fixed effect models generally preferable over random effects, if estimation of causal parameters from observational panel data is the goal (Halaby, 2004, 526; Allison, 1994, 181; Nickell, 1981, 1418).[6]

Fixed effects for choice data: The conditional likelihood approach Eliminating or estimating fixed effects is fairly straightforward in the linear model. Assume a simple linear model $y_{it} = X_{it}\beta + \alpha_i + \epsilon_{it}$,[7] where α_i is the unobserved heterogeneity. As is well known, given only two time periods ($t \in (1, 2)$), differencing across the two time periods eliminates the unobserved heterogeneity. This is the Difference-in-Difference estimator, $\Delta y_i = \Delta X_i \beta + \Delta \epsilon_i$, that is free of α_i, and can be estimated with standard OLS. Alternatively, α_i can be consistently estimated for panel data with more time periods simply by including unit dummy variables.

Voting behavior in the sense of electoral participation or party choice is generally measured as nominal choice data. Unfortunately, these fixed effect approaches do not work in non-linear models for discrete dependent variables such as logit or probit (Greene, 2004). First of all, there is no simple Difference-in-Difference estimator for non-linear models. Furthermore, consistent estimation of the α_i's is not possible if the number of time periods is small relative to the number of observation, as is commonly the case with social science panel data. This commonly known as the *incidental parameter problem* (Neyman and Scott, 1948). It can be shown that if α_i's are estimated with a dummy estimator, this leads to substantial bias in β estimates. β is estimated as 2β for two panel waves, and bias becomes negligible only with more than 16 panel waves (Katz, 2001; Abrevaya, 1997; see also Andersen, 1970). However, Chamberlain (1980), building on Rasch (1960) and Andersen (1973), showed that there is a method for logit models that is analogous to first-differencing in the linear model. The approach is generally referred to the conditional likelihood, the resulting models are the (binary) conditional logit[8], often also called the panel conditional logit (PCL) or fixed effect logit. The approach readily extends to ordered and multinomial choices as well - the following exposition of the conditional likelihood approach shall however only discuss the binary case.

Let y_{it} be a binary outcome, where 1 indicates a success for individual i ($i =$

---

[6]If causal identification is of secondary importance, choosing between fixed and random effects involves a bias-variance trade-off. See Clark and Linzer (2015) for a discussion and additional references.

[7]Let the continuous variable y at time t be the function of time-varying independent variables X with effect vector β, individual-specific time-constant parameters α_i and error term $epsilon_{it}$.

[8]Not to be confused with the closely related McFadden (1974) conditional logit, which is a generalized form of the multinomial logit model to alternative-specific independent variables.

$1, ..., N$) in time period t ($t = 1, ..., T$), 0 otherwise. Starting from a pooled conventional binary logit model where choice probabilities are a function of time-varying explanatory variables x_{it} and individual-specific time-constant "intercepts" α_i, the unobserved heterogeneity,

$$Pr(y_{it} = 1) = \frac{exp(x_{it}\beta + \alpha_i)}{1 + exp(x_{it}\beta + \alpha_i)}, \tag{3.1}$$

it can be shown that after conditioning on a sufficient statistic $s_i = \sum_t y_{it}$, the number of successes of individual i over all time-periods, the conditional probabilities are free of unobserved heterogeneity (Chamberlain, 1980). The resulting conditional probability of observing a particular choice sequence $Y_i = (y_{i1}, ..., y_{iT})$ given s_i is

$$Pr(Y_i|s_i) = \frac{exp(y_{it}x_{it}\beta)}{\sum_{\sum_t \lambda_t = \sum_t y_{it}} exp(\sum_t \lambda_t x_{it}\beta)}. \tag{3.2}$$

$\lambda_t(\lambda_1, ..., \lambda_{T)}$ denotes the set of all possible combinations of 0's and 1's. $\sum_{\sum_t \lambda_t = \sum_t y_{it}}$ is then the sum over all combinations of 0's and 1's such that the sum of the 1's in the combination equal the sufficient statistic. An example might come in handy here. Say T=3 and $s_i = 1$, i.e. one success was observed in three time-periods. There are three possible combinations that satisfy $s_i = 1$: (1,0,0), (0,1,0) and (0,0,1). Accordingly, the denominator becomes $exp(x_{i1}\beta) + exp(x_{i2}\beta) + exp(x_{i3}\beta)$. The conditional probability of observing $Y_i = (1, 0, 0)$ is $exp(x_{i1}\beta)$ divided by the denominator, $Y_i = (0, 1, 0)$ $exp(x_{i2}\beta)$ divided by the denominator etc.

The loglikelihood function is the sum of all realized choice sequences,

$$logL = \sum_n \frac{exp(\sum_t y_{it}x_{it}\beta)}{\sum_{\sum_t \lambda_t = \sum_t y_{it}} exp(\sum_t \lambda_t x_{it}\beta)}. \tag{3.3}$$

Note that conditional probabilities and loglikelihood are free of unobserved heterogeneity. As the log-likelihood function can be maximized independent of the α_i's, consistent estimation of β with unobserved heterogeneity being related in any way to the included variables is possible. Panel conditional logit is straightforward to implement and estimate since the structure of the conditional probabilities strongly resembles the conventional multinomial logit[9], and can in fact be estimated using a McFadden's conditional logit, a generalized version of the multinomial logit. The interpretation of coefficients also remains the same as in the multinomial logit model.

An important feature of the conditional likelihood approach needs to be pointed out here explicitly. Only information from individuals that change their voting

---

[9]For larger T, the denominator becomes unwieldy to program and slow to evaluate. Krailo and Pike (1984) noted that there is a recursive formula for the binary conditional logit that quickly generates the denominator for all Ts.

behavior between elections is used. If $s_i = 0$ or $s_i = T$, (for T=3, $Y_i = (0, 0, 0)$ or $Y_i = (1, 1, 1)$), there is only one combination that could have generated this outcome, and nominator and denominator are the same - the loglikelihood contribution of i is $ln(1) = 0$. This means that individuals with no variation on the dependent variable do not contribute to the estimation of β. The aggregate choice pattern is determined solely by the time-constant factors, β is only identified by the variation within individuals.

That only some of the available information is used is a frequent point of criticism of fixed-effect models in general. Fixed-effect models are inefficient in the sense that they disregard a substantial amount of variation in the data - all the between-unit variation. However, I think that this criticism misses the point. The particular purpose of the fixed effects approach is exactly to focus only on within-unit variation, with units of observation acting as their own controls. As Halaby (2004, 523) nicely puts it,

> the luxury of "throwing out" between variation is the very source of the advantage that panel data provide over cross-sectional data and that within-group estimators, such as fixed effects and first difference, exploit to avoid heterogeneity bias. Throwing out between variation is not wasting data: It buys protection against biased and inconsistent parameter estimates.

Related to the general issue of inefficiency is a particular concern that seems to come up frequently in discussions of the conditional likelihood approach. Only the units of observation that exhibit change in the dependent variable enter the estimation. Observations with no variation on the dependent variable effectively drop out. At first sight, this might raise concerns about problems of sample selection, i.e., selecting on the dependent variable. This is however fallacious (Brüderl, 2010, 987). It is simply that "such data points contain no information for estimating β, and so they should drop out of estimation" (Wooldridge, 2010, 621). That no change is observed is fully explained by the unobserved heterogeneity. For example, the behavior of a voter who participates in each election can be completely explained by setting α_i to infinity. If α_i may be infinitely high, there is nothing left to explain by the time-varying characteristics: The large impact of unobserved factors makes it impossible for the included variables to have any effect in these cases.

Although the exclusion of cases that have no variation on the dependent variable is not directly a problem of sample selection, the exclusion still has implications for the interpretation of the estimated parameters. This concerns the distribution of treatment effect magnitudes in the population. If treatment effects are homogeneous, i.e., the coefficient is the same for each individual, the exclusion will not bias the results. However, if treatment effects are heterogeneous, i.e., different subgroups have different coefficients, then the exclusion can weight the results to-

wards one subgroup.[10] This does not damage the integrity of the estimates for that subgroup, but should make us careful about generalizing the estimates to the whole population. The effect estimated by fixed effect models is always an Average Treatment Effect on the Treated (ATET), and not an Average Treatment Effect (ATE) (Brüderl, 2010).

Additionally, it is often brought up that a disadvantage of the fixed effects approach is that characteristics that are constant over time can not be included in the model. While it is of course true that all time-constant factors drop out, it is not true that time-constant variables cannot be included. This misunderstanding arises from failing to distinguish between time-constant effects and time-constant variables. "Within-group estimators sweep away only those time-invariant regressors with time-invariant parameters. Time-invariant variables with time-varying parameters are easily handled because neither within transformation eliminates them entirely." (Halaby, 2004, 525; see also Lee, 2012; Lee, 2015). If the effect parameter of a variable is allowed to vary over time, the variable can be time-constant, or time-varying. Time-varying parameters can be fairly easily estimated even with standard statistical software by specifying an interaction effect of the time-constant variable with time-period dummies. This is an important point especially in the context of this dissertation. The hypotheses I seek to test are specifically about variation in the parameters across elections.

While some of the criticisms about fixed effect models is based on an incomplete understanding of the method, there is one key disadvantage of the conditional likelihood approach. Since it effectively "gets rid" of the α_is, the unobserved heterogeneity, one looses the ability to evaluate individual choice probabilities. Thereby, one cannot calculate, say, average marginal effects. Depending on what the quantities of interest for a specific research undertaking are, the conditional likelihood approach may not be able to offer a satisfactory answer. However, this is not the case in the context of this dissertation. The quantity of interest that I seek to estimate is the difference in the parameters of interest across elections, which can of course be inferred directly from obtained parameter estimates.

Overall, I think that the conditional likelihood approach offers a very elegant analytical solution to the unobserved heterogeneity problem. It's key advantage is that parameters of interest are consistently estimated even if unobserved heterogeneity is related in any way to the included variables. This allows to improve inferences about parameters of interest without relying on the strong assumption of correct specification or strict exogeneity of unobserved or omitted variables. Its analytical approach means that two panel waves are sufficient to account for unobserved heterogeneity. This makes panel conditional logit especially useful in the field of electoral behavior research, where normally only very short panel designs are available.

In the above, I outlined my overall strategy to estimate the quantity of interest,

---

[10]I thank Prof Paul D. Allison for pointing this out to me in an email.

the difference in effect sizes across electoral arenas, from panel survey data. The theoretical argument does not only imply that EU-related policy preferences have a larger effect on voting behavior in EP elections when compared to FOE, but also compared to other second-order elections as well. In order to maximize leverage by maximizing the number of observable implications (King et al., 1994), the above discussed research design should therefore be ideally extended to include other elections as well. Fortunately, this is possible with the data that I will be using. As will be discussed in the next section that will introduce the data, voting behavior can be observed in three electoral arenas in the data at hand: In an EP election, a FOE and a state election, which qualifies as a different kind of SOE.

3.3 Case Selection: Multi-level elections in Bavaria 2013-2014

As has been pointed out above, panel data on individual preferences and voting behavior across multiple elections taking place a different levels of government is needed in order to enable strong inferences. This requirement alone drastically minimizes the set of suitable secondary data sets: The number of panel studies that cover elections at different levels of government, such as EP and FOE elections, and include measures of policy preferences, are rare. Moreover, in order to compare vote choices across multiple electoral arenas, the institutional context and the set of parties competing in these elections should ideally be held constant. Lastly, time is of the essence. If elections are too far apart, the potential interference from disturbances such as political or economic events becomes more likely. Finding a set of elections that satisfy these conditions, and having these elections covered by a good panel study is unlikely. In this light, a study that was designed specifically for the study of multi-level phenomena, the Making Electoral Democracy Work (MEDW) Bavaria Panel Study that I will use in this dissertation, comes surprisingly close to the ideal case. The study covers three legislative elections in the German state of Bavaria - State, Federal and European - that were held in relatively close succession in the years 2013 an 2014. In the following, I introduce the electoral context in which these elections took place, highlight factors that can be held constant by case selection and discuss how factors that vary between the elections can be accounted for at the modeling stage.

Institutional context Both the electoral system employed in the Bavarian state and in the German federal election is a mixed-member proportional representation (MMPR) system consisting of a first-past-the-post district tier and a proportional party-list tier. Both systems are fully compensatory, meaning disproportionality in the plurality tier is compensated via the PR tier (Shugart, 2001). It is therefore the party-list PR tier that determines the distribution of seats in the parliament. The plurality tier vote is far less important since it only determines which district

candidate represents the district in parliament. The main difference between the systems is that a special kind of open list MMPR is used in Bavaria (Massicotte, 2011), whereas lists are closed in Federal elections. For the purposes of studying party vote choice, the consequences of this difference for vote choice are negligible since voters make only limited use of the full possibilities offered by open list (Faas and Schoen, 2006; Rudolph and Däubler, 2016). European Parliamentary elections in Germany are conducted under (single tier) proportional representation with closed lists.[11] Even though there is variation in the kind of electoral system used at the different levels of government, I think there is a good point to be made that the central, decisive part of the electoral systems is the same. In State, Federal and EP elections, it is the proportional party list tier that preeminently defines electoral outcomes.

While the defining parts of the electoral system are relatively constant across elections, the institutional settings vary somewhat in the incentives for strategic voting. In the German context, two kinds of strategic voting incentives are generally distinguished: The wasted vote logic, and coalition (assurance) voting (Gschwend, 2007). The institutional factor that drives the wasted vote logic in the German PR system is the electoral threshold. Votes for a party whose vote share falls below the threshold will not obtain representation in the parliament, an outcome strategic voters seek to avoid. The threshold is set at 5% of the PR tier vote share in both the Bavarian state and Federal election. Until and including the 2009 EP elections, a 5% threshold was also employed in the EP election, but was reduced to 3% as part of an electoral reform in 2013. In February 2014, the German Constitutional Court overruled the 3% threshold in EP elections as unconstitutional, which meant that the 2014 EP election was conducted without any threshold in place.[12] The absence of a threshold in the EP election makes voting for small parties more viable than in State and Federal elections. Therefore, one should expect more sincere, policy-based voting in the EP election. The chosen modeling approach will allow to take the changing viability of parties into accout by allowing for election-specific intercepts. Moreover, the potentially increased overall role of policy considerations for vote choice in EP elections are not problematic, since I do not base my inferences on the absolute weight, but the relative weight of the two policy dimensions. Coalition assurance voting, in which supporters of the senior coalition partner "rent out" their list vote to the prospective junior partner if he is in danger of falling below the threshold, is also not a major concern (Gschwend et al., 2016). While some renting of votes from CSU supporters to the FDP in the State and Federal election, and not in the EP

---

[11]One peculiarity of the EP election system in Germany is that parties can choose between running a nation-wide or state-specific list in the EP election. The nature of the list is inconsequential as state-lists are merely used when seats that are obtained by a party nation-wide are distributed to the lower level. In 2014, only the CDU/CSU used state-specific lists which is primarily due to the state-specific character of the CSU.

[12]As 96 German MEPs were elected in a single national-level district, the effective threshold was at around 0.7% (Lijphart, 1994; Taagepera, 2002; Gallagher and Mitchell, 2005).

election, might have happended (Stoetzer et al., 2015), these two parties take very similar positions on both Left-Right and EU integration issues, and will therefore not distort the policy-based voting rationale.

Political and party system context   The Bavarian multi-party system is heavily dominated by the conservative, Christian-Democratic CSU (*Christlich-Soziale Union*). On the State level, the CSU has been in government since 1946 and often forms a single party government. Acting as the sister-party of the CDU, the CSU also participates in the Federal government coalition since 2005. The social-democratic SPD (*Sozialdemokratische Partei Deutschlands*) is the largest opposition party in Bavaria. After the 2013 Federal election, the SPD joined a government coalition with the CDU/CSU at the federal level. The free market-liberal FDP (*Freie Demokratische Partei*), coalition-partner to the CSU at the state and the CDU/CSU at the federal level, lost parliamentary representation in 2013 at both the Bavarian state and Federal level by falling below the electoral threshold of 5%. Four more parties make up the opposition, of which two, the Free Voters and the newly founded AfD, merit further introduction. Traditionally rooted at the local level in Bavaria, the Free Voters (*Freie Wähler*) competed for the first time in 1998 at the State level, and could realize considerable vote shares in State elections in recent years. Although the Free Voters are relatively heterogeneous in their political profile, about two-third of the voters tend to be recruited from CSU-leaning voters (Kießling, 2008). The right-wing populist Eurosceptic AfD (*Alternative für Deutschland*) was founded in February 2013 in opposition to the Euro crisis and Euro country "bailouts" (Arzheimer, 2015). The AfD did not run in the Bavarian State election, but in the Federal and 2014 EP election.

The period between September 2013 and June 2013, in which the three elections under investigation were held, can be described as politically quiet. Apart from the elections that took place and their results, no major events or developments can be identified that are likely to have influenced the political landscape. Even though the overall political context was very much stable, there is substantial variation in the party fortunes between the elections, as shown by Table 3.3, that displays the official election results in the State of Bavaria. Two important patterns can be identified. First of all, the party results that vary substantially between State and Federal elections are those of the Free Voters, and to a lesser degree those of the CSU, FDP and Left. As the organization of the Free Voters is heavily focused only on the State of Bavaria, and structurally weak in all other States, it is generally not perceived as a viable option in Federal elections, and gets strategically deserted in Federal elections. The CSU benefited disproportionately from the votes of former Free Voters voters. For the Left, it is the other way around - it is generally not viable on the State level, but is safe to pass the electoral threshold at Federal elections. The second pattern in the official results is that the CSU lost substantially in the EP election, while the Greens and AfD gained substantially.

First of all, the variation of party fortunes may be regarded as a positive feature.

	State	Federal	EP
CSU	47.7	49.3	40.5
SPD	20.6	20.0	20.1
Greens	8.6	8.4	12.1
Free Voters	9.0	2.7	4.3
FDP	3.3	5.1	3.1
Left	2.1	3.8	2.9
Pirates	2.0	1.9	1.2
AfD	–	4.3	8.1

Table 3.1: Official Election Results: Party vote shares

It indicates that there is variation in the vote choice of Bavarian citizens, variation that needs to be explained. If there were no variation, there would be nothing to explain. However, some of the variation of party fortunes across the elections may also be problematic, as it may potentially bias our inferences. We can never be sure whether the variation is the results of changes in the level of government, or merely the consequence of time. Fortunately, the variation of party fortunes over time can be easily adjusted for at the modeling stage by allowing for intercepts that vary across elections. To help the readers' intuition - if we were to estimate a model of vote choice only with the intercepts on the right-hand side, the point estimates would describe the changes in the observed vote shares of the parties between the elections, i.e., the predictions would generate Table 3.3. Empirically, this encompasses variation in institutional context that favors particular parties in particular electoral arenas, or factors located at the party level. Put differently, the intercepts capture all institutional and party-level factors that play out equally in the vote choice calculus of all respondents. It is to be noted that the intercepts therefore capture implicitly many aggregate-level explanations put forward by second-order election theorists to describe the variation of party fortunes between first- and second-order elections - such as swings in the overall party popularities between elections, whether a party is a government or opposition party, or a small party that benefits from the absence of strategic voting incentives in second-order elections, or a conservative/green/socialist/Eurosceptic party.

To summarize, the case of the 2013-2014 Bavarian multi-level election comes as close as is possible to the ideal case for a study of multi-level voting behavior in the real world. The electoral systems are very much comparable over the elections, and the close timely proximity of the election and the stable party system and political context allow to credibly exclude the possibility that changes in voting behavior are merely the consequence of changes over time, rather than the change in the level of government. Two problems are identified: The party fortunes vary substantially across the elections as a consequence of strategic voting considerations. However, as it turns out, these variations should be easily controlled for in statistical models. Moreover, one party, the AfD, did not run in one election under

investigation. As will be explained below, this problem can be tackled by using additional measures available in the data at hand. Overall, the case of the 2013-2014 Bavarian multi-level election constitutes a most promising empirical case to study multi-level voting behavior.

3.4   The MEDW Bavaria Voter Panel study

The data I am analyzing is part of the Making Electoral Democracy Work project (MEDW) (Blais, 2010). The panel covers the Bavarian State and German Federal elections in September 2013, as well as the 2014 European election and was purposefully designed to study voting behavior of the same individuals in multiple electoral arenas. The MEDW Bavaria Panel was administered by the polling firm Harris Decima in cooperation with Infratest Dimap, and conducted as an online survey for which respondents were recruited offline. Offline recruitment ensured that, unlike other online surveys, the MEDW panel does not suffer from systematic overrepresentation of young and/or left-leaning voters. Respondents were surveyed in the week before and after each election, all in all five times.[13] Good quality control and attractive compensation schemes lead to an unusually low panel attrition rate, which meant that the target sample size of 1000 respondents in the last wave was substantially surpassed. In the end, 2854 respondents participated in all panel waves - without panel refreshments.[14]

Measuring voting behavior   A panel participant is counted as having participated in an election if he/she indicated to already have cast an absentee vote in the pre-election wave, or if he/she indicated to have voted in the post-election wave. 302 participants refused to indicate whether they turned out or not in at least one of the elections - I drop these from the data. Participants who reported to have turned out where asked to indicate their party vote choice. Two of the three elections took place under a two-tier electoral system. As discussed in detail above, I only use the party-list PR tier votes as measures of party vote choice. The party set encompassed the CSU, SPD, Greens, Free Voters, FDP, Left, Pirate Party and AfD. Respondents received the same set of parties in the same order in each survey wave, with the exception of the AfD in the State election, in which the party did not compete. Respondents could also indicate "Other party", "Wasted vote" and "Don't know" - these I treat as missing information. In my analyses of party vote choice, I confine my sample to panel participants who voted for one of the eight parties in all elections. There were 1896 participants who reported to

---

[13]Since the Federal election took place only one week after the State election, the second panel wave was a combined post-election survey for the State election and pre-election survey for the Federal election.

[14]For further information on the data set visit electoraldemocracy.com/voter-behaviour or see Golder et al. (2017).

have turned out in all elections. Of those, 527 did not indicate a party vote for one of the eight parties in the party set. The vast majority of these chose "Don't know", therefore indicating either not remembering party vote choice or refusing to report party vote choice. Thereby, 1369 remaining panel participants constitute my core sample for the analysis of party vote choice. I discuss the selectivity of the sample and employ alternative operationalizations, mainly for robustness testing, in the relevant Chapters 5,6,7.

Table 3.2 investigates the representativeness of the sample by comparing the observed turnout and party vote choice rates with the official results of the three elections in the state of Bavaria. To make official party results comparable, they are transformed into the percentage of votes the parties received as a share of the votes cast for any of the eight parties included. Furthermore, the difference between sample percentages and official percentages is calculated.

	Turnout	CSU	SPD	Greens	FV	FDP	Left	Pirates	AfD
State									
Sample	92.9	45.7	24	10.1	9.8	5.4	2.4	2.6	.
Official	63.6	51.1	22.1	9.2	9.6	3.5	2.3	2.1	.
Difference	(29.3)	(-5.4)	(2)	(0.9)	(0.1)	(1.9)	(0.2)	(0.4)	.
Federal									
Sample	95.3	46.4	23	8.7	3.5	5.6	4.5	1.9	6.4
Official	70	51.6	20.9	8.8	2.8	5.3	4	2	4.5
Difference	(25.3)	(-5.2)	(2.1)	(-0.1)	(0.7)	(0.2)	(0.5)	(-0.1)	(1.9)
EP									
Sample	77.1	40.2	22.4	11.9	6.1	3.1	4	1.5	10.9
Official	40.9	43.9	21.8	13.1	4.7	3.4	3.1	1.3	8.8
Difference	(36.2)	(-3.6)	(0.6)	(-1.2)	(1.4)	(-0.3)	(0.9)	(0.2)	(2.1)

Table 3.2: Comparison of sample turnout and party choice rates with official election results

Table 3.2 offers two important insights: First, there is a considerable difference between turnout rates observed in the sample and official turnout in all three elections. Observed turnout varies between 93% (+29%) in the State, 95% (+25%) in the Federal and 77 (+36%) in the EP election.[15] Such discrepancies between observed sample and population turnout is not specific to this particular data set, but are a phenomenon observed in virtually all survey studies. There are two main explanations for this: overrepresentation of actual voters in the sample, and overre-

---

[15]The level of turnout overreporting in the data seems higher than usual (e.g. Schmitt and Mannheimer, 1991; Swaddle and Heath, 1989). There are two plausible explanations: Firstly, non-response bias is higher in panel studies as panel attrition tends to be stronger among respondents with low turnout probabilities. Secondly, respondents might alter their behavior simply because they are a participant in a study (Hawthorne effect), which might even be more pronounced in a panel study (Selb and Munzert, 2013, 192). In the present panel data, the effect of being repeatedly contacted by a polling firm over the course of almost one year might have a cumulative effect on electoral participation.

porting by actual nonvoters (Selb and Munzert, 2013). Important methodological issues arise from these that will be addressed in detail in my analysis of turnout in Chapter 4. I also show how panel data can be leveraged to buy some protection against these issues. Second, looking at the party vote choices, the partisan balance of the sample seems satisfactory. CSU voters are underrepresentated, and SPD and AfD voters are slightly overrepresented in the sample, compared to the official election results. However, I would suggest that these deviations are minor and overall partisan balance surprisingly good, given panel attrition effects and the necessarily selective construction of my party vote choice sample.

Measuring policy preferences Policy preferences are measured using simple 11-point policy scales, on which panel participants place themselves and the parties multiple times over the course of the panel study. I deliberately choose this very basic and conventional operationalization strategy for policy preferences in order to keep my analysis comparable with the literature I am contributing to. The vast majority of EP election studies have used the EES Voter Survey. The MEDW Bavaria Panel used the exact same phrasing of the question on European integration as in German version of the EES (see e.g., Schmitt et al., 2016). Voter opinion on European Integration are measured by asking respondents to locate themselves on a 0-10 11pt scale ranging from "Integration has gone too far" to "Integration should be pushed further."[16] Directly after this item, respondents In the MEDW Bavaria panel were also asked to place all eight parties on the same scale. Both items were asked two times throughout the panel study, early on in the pre-State election wave in September 2013, and again in the June 2014 pre-EP election wave. Left-Right self-placements and party positions are measured analogously: Respondents were asked to place themselves on a 0-10 11pt scale[17], ranging from "Left" to "Right", and then then asked to locate the eight parties on the same scale. The left-right self-placement and party measures were part of the MEDW panel in each pre-election wave, i.e., three times.

To infer European integration preferences for the Federal election, I use the measures from the pre-State election panel wave as a proxy. I think this is unproblematic since the Federal election took place only one week after the State election, and respondents are unlikely to have changed their views on Europe in such a short time frame. Additionally, I reduce missing information by inserting the most timely measure of policy self-placement that is available for the respondent. E.g.,

---

[16]"Some say European unification should be pushed further. Others say it already has gone too far. What is your opinion? Please indicate your views using a scale from 0 to 10, where '0' means unification"has already gone too far" and '10' means it "should be pushed further"." Party scales were introduced with "Where would you place each of the following political parties in Bavaria on that scale?", with the eight parties being arranged in a item battery.

[17]The question was "In politics people sometimes talk of left and right. Where would you place yourself on a 0 to 10 scale where 0 means far left and 10 means far right?". Party scales were introduced with "Where would you place each of the following political parties in Bavaria on that scale?", with the eight parties being arranged in a item battery.

if a respondent refused to indicate a self-placement in the pre-EP election wave, I insert the self-placement from the pre-Federal election wave instead if available. I think this is the best proxy for a respondents EU integration placement that is available, and helps to eliminate potential selection bias due to missing data. I apply the same procedure to Left-Right placements. Missing party placements are not inserted.

Measuring party positions I operationalize party positions as the mean party placements across all respondents, separately for each election. This strategy is most commonly used and considered "best practice" to obtain objective party positions because it buys protection against projection bias that is inherent in subjective, individually perceived party positions (Conover and Feldman, 1982; Granberg and Holmberg, 1986; Merrill III et al., 2001; Adams et al., 2005; Grand and Tiemann, 2013). Moreover, it avoids scalability issues that would arise if party positions were inferred from external sources, such as expert surveys or party manifestos.

Figure 3.4 displays the estimated party positions in the two-dimensional European policy space. For clarity's sake, I only display the party position in the Federal election and EP election.[18] As a frame of reference, I additionally plot the party positions from the 2014 Chapel Hill expert survey (Bakker et al., 2015). First of all, movement of the mean perceived party positions between Federal and EP election is minor. Party positions seem more volatile on EU integration than on Left-Right. Most parties seems to be perceived to move slightly to the center on EU integration before the EP election, but are largely constant on the Left-Right dimension. The AfD is an exception to this pattern, as it is perceived distinguishably more rightist and slightly more anti-EU at the EP election. Comparing the mean perceived party placements to the party positions estimated by the Chapel Hill expert survey, one can see surprisingly little differences. For all parties except the AfD, survey respondents and experts seem to largely agree about the party positions on both dimensions. The difference between voters and experts about the Left-Right position of the AfD is remarkable. One plausible explanation may be that the AfD was founded primarily as a Eurosceptic party. The strong emphasis on European issues (voters and experts seem to agree about the AfD position on this dimension) might make it difficult for voters to place the newly-founded AfD on Left-Right issues, on which the party had largely remained silent (Arzheimer, 2015; Jankowski et al., 2016).

---

[18]Party positions across the State and Federal election, which took place only one week apart, are practically indistinguishable.
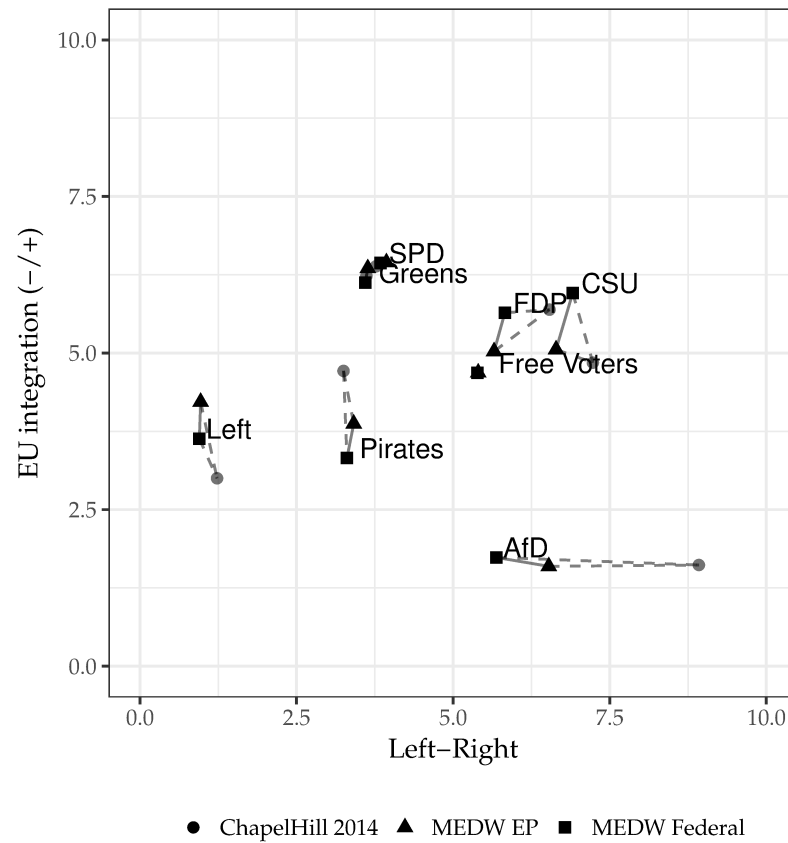
Figure 3.4: Mean perceived party positions in 2013 Federal and 2014 EP election (movements in solid lines), and 2014 ChapelHill expert survey (in dashed lines). For Free Voters, ChapelHill not available.
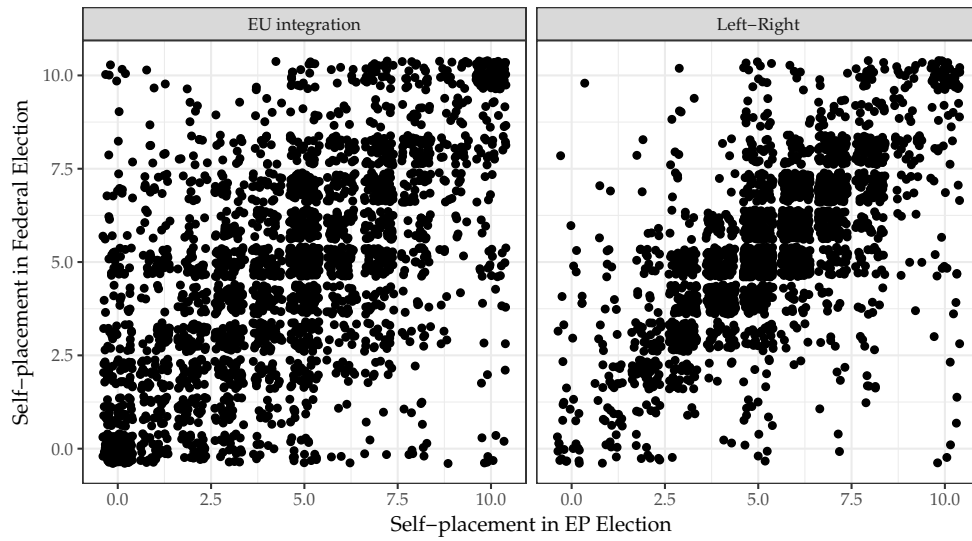
Figure 3.5: Jittered voter self-placements on Left-Right across Federal and EP election. Correlation coefficients .6 and .67
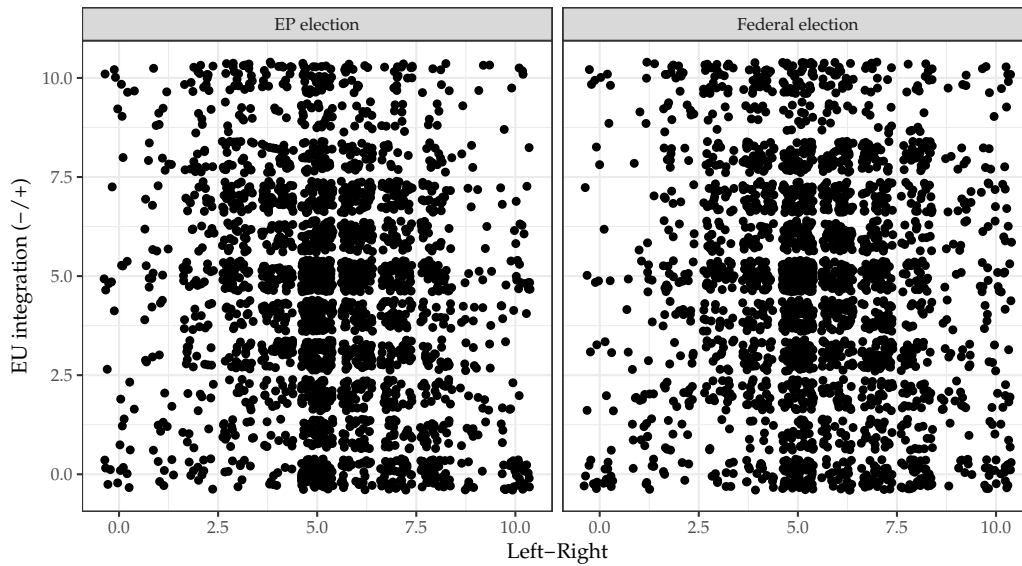
Figure 3.6: Jittered self-placements of EU integration and Left-Right. Correlation coefficients -.06 and -.16

Chapter 4

Legitimacy or Alienation? EU integration preferences and differential turnout

4.1 Introduction

Does Euroskepticism drive low turnout in European Parliamentary (EP) elections? With EP election turnout rates at an all-time low and a Europe-wide trend towards rising Euroskepticism in years of political and economic crisis, it is vital for policy makers and researchers to understand whether there is a causal relationship between the two. While the conventional wisdom of the public falsely tends to treat the existence of such a relationship as a self-evident truth, scholars of EP elections remain divided. Political scientists are currently not in a position to base their advise to the public on a broad academic consensus. Clearly this is not a satisfactory state of European election research.

Divisions among scholars persist due to conflicting theoretical backgrounds. Second-order theory (Reif and Schmitt, 1980), the traditional theoretical lens through which European elections have been analyzed, argues that the preferences and attitudes of citizens about the EU and its policies do not play a decisive role for voting behavior in European elections. Even though the election is supposed to be about Europe, citizens are said to behave just like in a less important election at the national level. This transfer hypothesis (Clark and Rohrschneider, 2009), that citizens apply their *domestic* decision-making calculus to the European level, has over the years become the essence of second-order thinking. In the recent decade a body of research has emerged that particularly challenges the validity of the transfer hypothesis. The *Europe matters* argument centers on the proposition that voting behavior in EP elections is at least partially driven by considerations that are specific to the European arena. These considerations entail attributions of legitimacy and policy preferences over the extent of European integration that

range from opposition to support for EU institutions and its policies (see de Vries and Hobolt, 2016, for an overview).

While the theoretical points of contention have been worked out very precisely, mixed empirical evidence from often weak research designs have so far prohibited the field from unifying behind a consensus position. Empirical studies from second-order theorists generally find no indication that preferences and attitudes towards European Integration have a non-negligible influence on EP turnout (Schmitt, 2005; Schmitt and van der Eijk, 2007; Marsh, 1998; van der Eijk and Franklin, 1996), those of proponents of the Europe matters argument do (e.g. Stockemer, 2012; Hobolt et al., 2009; Hobolt and Spoon, 2012). This contribution will certainly not trigger such a consensus. But what I want to attempt here is to provide a more definite answer to the fundamental question how EU preferences and attitudes are connected to electoral participation in EP elections. Unlike previous research that has sought to provide a broad picture of the correlates of EP election participation across member states based on aggregate turnout rates (e.g. Stockemer, 2012; Mattila, 2003; Studlar et al., 2003) or individual-level voter survey data (e.g. Schmitt and van der Eijk, 2007; Hobolt et al., 2009; Stockemer, 2012),[1] my primary concern is the internal validity of the supposed relationship. Previous research designs often faced unsurmountable methodological challenges that undermined their inferential potential considerably. Since aggregate-level analyses are necessarily exposed to the danger of ecological fallacy, the more recent literature has converged on individual-level data from voter surveys, most commonly using the European Election Study (EES). Limited to cross-sectional survey data, previous research had to rely on retrospective measures of previous voting behavior, which opened the door for potential retrospective bias. Cross-sectional analyses are also notoriously prone to specification bias due to unobserved heterogeneity. Furthermore, the consequences of overreporting of voting for the validity of inferences were not addressed by previous contributions. In the face of these potentially multiplicative and/or countervening biases, it is no wonder that the academic verdict is still out whether European preferences influence electoral participation in EP elections.

The present study proposes a research design that comes to grips with some of the methodological obstacles that previous studies have faced. In the first part, I explore the three competing main theoretical arguments about the relation between European preferences and EP turnout and work out the testable implications. In a second part, I address methodological challenges and develop a research design that allows for strong inferences. The empirical analysis employs the MEDW Bavaria Panel which traces individual voting behavior across the 2013 State, the 2013 Federal and the 2014 EP election. I analyze the data using a panel conditional logit model that is particularly well suited for the analysis of short panels and allows to control for unobserved heterogeneity and individual-specific overreporting.

---

[1]A notable exception is Hobolt and Wittrock (2011), who employ experimental methods.

4.2 Theoretical background

Electoral participation is the subject of a vast literature (for an overview see Blais, 2006; Geys, 2006; Aldrich, 1993). The key factors for turnout can be separated into two categories: Individual-level factors and context-level factors. While individual factors primarily explain why some citizens are more likely to turn out than others, contextual factors explain why turnout levels are higher in some elections or countries than in others. The question whether European preferences drive turnout in EP elections operates at the intersection between individual- and context-level factors.

The EU preferences of a voter - broadly defined as a set of preferences about specific EU policies and integration efforts - are an individual-level characteristic. However, the research question is not concerned with the effect of EU preferences on the likelihood of voting in EP elections per se. It might be that EU preferences affect turnout in *any* election. Of course, this is not what is implied by the research question. Rather, what we want to know is whether EU preferences have a stronger effect on turnout in EP elections than in other elections. Stating it even more precisely, the research question is how the effect of EU preferences on electoral participation interacts with the change in the electoral context, that is the change of the electoral arena from the national arena to the European arena.

The second-order argument on EP turnout The traditional theoretical framework to analyze electoral outcomes between different electoral arenas is the second-order national elections model (Reif and Schmitt, 1980; Reif, 1997). Second-order theory is based on the idea that there is a hierarchical relationship between the electoral arenas that is defined by the importance that citizens assign to elections held in the arenas. Election in second-order arenas where there is "less at stake" are subordinate to elections in the first-order arena, which are defined as serving the function of electing the head of government. The subordinate role of second-order elections expresses itself in the way citizens think about and behave in these elections. Citizens are said to have no considerations or preferences that are specific to the second-order arena. Instead, they get their bearings from the first-order arena. This "transfer hypothesis" (Clark and Rohrschneider, 2009), that citizens simply transfer their first-order decision-making calculus to the second-order arena even if the election is about something else is "the essence of the second-order interpretation" (Marsh and Mikhaylov, 2010, 13). As follows from this line of reasoning, second-order theorists have argued that EP elections are not *European*, but (second-order) *national* elections. Even if this is the case, second-order theory does not preclude a potential influence of European issues on electoral behavior in EP elections. Second-order theory merely suggests that the issues remain the same even if the arena changes. So if EU preferences have an

effect on voting behavior in the national arena[2], the effect should be the same in the European arena.

A growing body of research has emerged in the recent decade that challenges the transfer hypothesis and argues for a role of EU preferences in explaining electoral participation in EP elections. Two different theoretical arguments can be distinguished that have emerged in the literature. First, an explanation that is based on the perceived legitimacy of the European Union and second, an explanation based on policy alienation on European integration issues (see Schäfer and Debus, 2015, for this useful typology).

The legitimacy argument on EP turnout Legitimacy-based studies (Stockemer, 2012; Flickinger and Studlar, 2007; Mattila, 2003; Studlar et al., 2003; Blondel et al., 1997) have not explicitly specified the causal mechanism that relates EU preferences to differential turnout in EP elections. These studies merely state that pro-EU integration preferences should increase a citizen's likelihood to turn out, anti-EU integration preferences should decrease turnout likelihood. Implicitly, the argument is based on the proposition that citizens turn out based on the degree to which they support the political system. Taking cues from the Voice-Exit literature (Hirschman, 1970; Weber, 2011), withholding participation is a means to discredit a political system that is seen as illegitimate, while participation yields psychological rewards for citizens with strong system support. The crucial point that the argument makes is that citizens assign specific legitimacies to specific electoral arenas. In order to explain differential participation across different electoral arenas, attitudes towards the EU need to have a larger effect in EP elections than in national elections. Moreover, the argument has clear implications about the functional form of the relationship between EU preferences and turnout. If the legitimacy story is correct, an increase in EU support should always lead to more participation. One should therefore expect a monotonic relationship between EU preferences and the likelihood to turn out in EP elections.

The policy alienation argument on EP turnout Studies arguing based on a EU issue voting model (Clark and Rohrschneider, 2009; Hobolt et al., 2009; de Vries et al., 2011; Hobolt and Spoon, 2012) have been much more specific about the theoretical foundations of the relationship between EU preferences and turnout in EP elections. At the heart of the argument is a model of multidimensional spatial voting. European citizens are said to be located in a two-dimensional policy space consisting of a Left-Right and EU integration dimension. While the Left-Right dimension represents the traditional line of national party-competition, the EU integration dimension captures diverging preferences over the extent of European integration that are cross-cutting left-right issues. As discussed in Chapter 2, the

---

[2]de Vries and Hobolt (2012), de Vries (2010) and de Vries (2007) have shown that EU issues matter at the national level in many elections.

argument hinges on the proposition that citizens assign different weights to the issue dimensions in different elections - they "recalibrate their issue space" (Marsh and Mikhaylov, 2010, 13). They are said to do so by altering the relative salience of the EU integration dimension. I.e., policy distance on the EU integration dimension becomes more salient relative to policy distance on the left-right dimension to voters in European arena than in national arenas.[3]

How does this tie in with voter turnout? If voters care more about the positions parties take on European integration in EP elections than in national elections, this might lead to policy alienation from the parties citizens usually vote for in the national arena. Moreover, whether EU preferences lead to policy alienation is fundamentally contingent upon party system characteristics. Voters can only feel alienated if the party system does not offer political alternatives that represent voters' recalibrated policy preferences. Ultimately this means that the relation between EU preferences and differential turnout might vary a great deal between different party-systems and electoral contexts. While empirical research clearly needs to carefully account for the features of party systems, this does not absolve me from proposing a general model. The most widely accepted model of party-competition in the two-dimensional European policy space among European election scholars is the inverted U curve (Hooghe et al., 2002). It states that left-right centrist catch-all parties tend to adopt similarly pro-European positions, while left-wing and right-wing parties, even if for different reasons, take Euroskeptic or anti-EU positions. Given such a typical constellation of party positions in the European policy space, voters at risk of high policy alienation in EP elections can be identified. Firstly, these are left-right centrist voters with anti-EU preferences. As all centrist parties are also pro-European, these voters would have to sacrifice too much policy closeness on the left-right dimension to find a matching anti-EU party. The second high-risk group are pro-European voters that hold extreme positions on the left-right dimension. These voters do not feel represented by leftist or rightist parties on European issues and would have to give up too much on left-right closeness in order to get their preferred position on European issues, since the political alternatives are located at the center of the left-right dimension.

This verbal investigation of the multidimensional spatial model shows that the increased policy alienation in EP elections, due to the recalibration of the issue space, does not only depend on EU preferences, but left-right preferences as well. Policy alienation on EU preferences is conditional on Left-Right preferences.[4] The voters that are most likely to be feel increasingly alienated in EP elections, and

---

[3]Whether European issues matter more in EP elections to voters is to be investigated in Chapter 5.

[4]In terms of a statistical model, such a relationship can be modeled with interaction terms, potentially with added polynomial terms to account for non-linearities. Previous research has not adequately taken into account the conditionality of the relationship. E.g., while Hobolt et al. (2009) control for the left-right position of respondents in their empirical analysis, there is no interaction effect with EU integration preferences to capture this prediction of the alienation argument.

are therefore more likely to abstain, are the voters with no political alternatives in their left-right vicinity that match the voters' EU preferences. In a typical party system, in which multidimensional party positions follow an inverted U-shaped, the conditional relationship should therefore follow a U curve. Following the left-right continuum, leftist voters with more pro-EU preferences, centrist voters with more anti-EU preferences, and rightist voters with more pro-EU preferences should feel increasingly alienated, and be less likely to turn out.

To summarize the three theoretical arguments: The second-order argument, the legitimacy argument and the policy alienation argument have diverging implications for the relationship between EU preferences and turnout in EP elections. These can be formulated as

H_0: *EU preferences have the same effect on turnout in EP elections as in first-order elections. (second-order model)*

H_1: *EU preferences have a larger effect on turnout in EP elections than in national elections. The likelihood to turn out increases monotonically with increasingly pro-EU preferences. (legitimacy model)*

H_2: *The effect of EU preferences on turnout in EP elections varies conditional on the left-right position of voters and the constellation of parties in the European policy space. (policy alienation)*

4.3 Research Design

Empirical Challenges

Retrospective partipation measures Studying voting behavior across the electoral arenas ideally requires panel survey data that tracks individual turnout decisions across multiple elections. While this insight is well appreciated throughout the literature, such data has until recently not been available. Instead, previous research designs have almost exclusively relied on retrospective measures of previous voting decisions of respondents. It is well established that retrospective measures of voting behavior come with systematic distortions (Benewick et al., 1969; Weir, 1975; Waldahl and Aardal, 2000; van der Eijk and Niemöller, 1979). Substantial proportions of respondents tend to mis-remember or mis-represent former voting decisions. Retrospective evaluations tend to err systematically in the direction that makes previous evaluation seem more consistent with current evaluations. These distortions have methodological implications (Shachar and Eckstein, 2007). The magnitude of retrospective bias seems to be often systematically related to key individual-specific characteristics that also relate to turnout. E.g., more educated citizens feel a higher pressure to appear consistent and are more likely to mis-represent their previous voting decisions in order to appear consistent. In the context of this investigation: If more educated citizens tend to have more pro-EU preferences, this will lead to confounded estimates of the effect of EU preferences

on electoral participation. Due to the multitude of potential causal relationships, it is unlikey that standard techniques of statistical control can credibly eliminate retrospective bias. In the most commonly used data set, the European Election Study (EES), respondents are asked at the time of the EP election whether they voted in the general national election or not. Here, retrospective bias is especially threatening, as respondents from different countries vary in the distance in time to the last first-order election, which should introduce systematic variation in the magnitude of retrospective bias, since the likelihood to mis-remember or mis-represent voting behavior seems to increase with time (Blank et al., 2003). Such retrospective bias risks can be substantially reduced with panel survey data that polls voting decisions simultaneously for all respondents only a couple of days after each voting decision.

Overreporting of turnout   Another common problem in the study of electoral participation on observational data is systematic overreporting of turnout. Two reasons for overreporting of electoral participation in survey data can be distinguished: the overrepresentation of voters among respondents (non-response bias), and the mis-reporting of actual abstainers as voters (measurement bias) (Selb and Munzert, 2013). While the former means that the sample is not representative of the wider electorate, the latter concerns the misclassification of actual voting behavior.

Non-response bias is generally considered to occur at the sample selection stage. The problem of non-response is notable aggravated in panel data, as panel attrition tends to be stronger among those with low ex-ante probabilities to turn out. Another prominent cause is the Hawthorne effect in election surveys (Clausen, 1968; Persson, 2014). Here, ex-ante probabilities to turn out are increased merely by participation in a (panel) study on voting behavior. From a causal inference perspective, non-response bias results from treatment effect heterogeneity, that is the average effect estimated among citizens who take part in the survey is (in expectation) different from the effect among all citizens. Non-response bias is therefore primarily a problem of external validity as it undermines our ability to generalize our inferences from a given sample to the population. While there are some options to deal with non-response bias (e.g., Heckman, 1979; Manksi, 1995; Jackman, 1999), the problem shall not be addressed in the following analysis, primarily because the effect estimated by panel fixed effect models are never average treatment effects anyways. Problems of generalizability are discussed in the concluding section.

Measurement bias due to misreporting is generally seen as much more dangerous for the validity of the inferences of empirical research.[5] Katz and Katz (2010) warn against disregarding the issue of misreporting as this might lead to "poten-

---

[5]Ansolabehere and Hersh (2012) find that "most of the overreporting of turnout is attributable to misreporting rather than to sample selection bias" even in Web-based surveys. Selb and Munzert (2013) arrive at a more balanced verdict.

tially distorting the relative impact of the characteristics of interest on the response variable and leading to erroneous conclusions". Neuhaus (1999) shows that even response misclassification that is independent of covariates leads to attenuation bias, i.e., coefficient estimates are biased towards the null hypothesis. However, if respondent-specific misclassification probabilities are dependent on covariates, which research comparing validated and reported votes has convincingly shown (Silver et al., 1986; Bernstein et al., 2001; Cassel, 2003), biases can go either way. In such situations, relatively small amounts of misclassification can lead to significant bias (Hausman et al., 1998). Seen from a causal inference perspective, bias is in this case due to confounding factors that are both related to turnout and misreporting. E.g., since educated citizens are more likely to overreport, estimates of the effect of education on voting are most likely overestimates if misreporting is not accounted for (Bernstein et al., 2001). Similar confounding factors might be at play for EU preferences. If those most supportive of the overall regime are more likely to overreport and also tend to hold more pro-EU preferences, this might lead to an overestimate of the effect of EU preferences on reported turnout. Previous research has not specifically addressed potential bias due to overreporting. This paper acknowledges that in the absence of validated voting records, voting decisions in survey data are reported and not actual voting decisions. As I will show, overreporting can be framed and modeled as a form of unobserved heterogeneity. Standard techniques such as adding control variables for potential confounders are unlikely to succeed in adequately addressing the fundamental problem of measurement bias. As will be outlined in detail in the next section, the proposed research design allows to partially account for overreporting by taking into account respondent-specific propensities to overreport.

Unobserved heterogeneity   Much of the turnout literature is dedicated to explaining which factors determine individuals' inclination to turn out. Among the most prominent factors are political interest, socio-economic factors such as income, status and education, or a general habit of (non-)voting that individuals have developed over the years, just to name a few. The important point about individual-specific effects is that they are not confined to one electoral arena, but determine an individual's innate ability to (report) voting in *any* election or electoral arena. In the econometric literature, individual-specific effects are referred to as incidental parameters or unobserved heterogeneity. Even though individual-specific effects are constant across arenas, this does not mean that they can be disregarded if one wants to explain differences in electoral participation across arenas. As has been discussed in Chapter 3, unlike in linear models, first differencing does not solve the incidental parameter problem in non-linear models such as Logit or Probit that are used to study electoral participation. Much of the research on differential turnout between national and EP elections has used the difference in individual turnout between one election and another as the dependent variable - with voters that turned out in one, but not the other, being coded as differential

voters (e.g. Schmitt, 2005; Hobolt et al., 2009). While this operationalization of the dependent variable might reflect a substantive interest in the motivations of these voters, from a statistical perspective it is not clear what the differencing should accomplish. As has been discussed in Chapter 3, not accounting for individual-specific effects opens the door for heterogeneity bias. Standard techniques such as including control variables are unlikely to succeed in reducing heterogeneity bias. Especially since individual-specific factors, such as an individual's inclination to misreport turnout behavior remains unobserved.

A panel conditional logit model of turnout in multiple elections

The probability of an individual i to report to have turned out in an election t, $P(y_{it} = 1)$ ($y_{it} \in 0, 1$), can be thought of as a function of the true likelihood to turn out in the election (η_{it}), plus the inclination to misrepresent turnout behavior (θ_{it}), i.e. $P(y_{it} = 1) = f(\eta_{it} + \theta_{it})$. η_{it} can be further separated into election-specific factors and individual-specific factors that are constant across elections, and a random error component: $\omega_{it} + \mu_i + \epsilon_{it}$.

The obvious difficulty here is the "cheating parameter" θ_{it} that represents the inclination to misrepresent turnout behavior. An individual's inclination to cheat of course remains unobserved. To come to terms with this problem, I make a simplifying assumption: That individuals' inclination to cheat are constant across elections, i.e. $\theta_{it} = \theta_i$. While this assumption is certainly a bold one, it is consistent with the literature on overreporting, which has to my knowledge only theorized about the connection between individual-specific factors and overreporting, and not about variation in overreporting across elections.

Having made this assumption, the reported turnout function can be simplified by collecting terms. $P(y_{it} = 1) = f(\omega_{it} + \alpha_i + \epsilon_{it})$, where α_i is the sum of the cheating parameter and the individual-specific turnout factor, i.e., $\alpha_i = \theta_i + \mu_i$. The election-specific effect ω_{it} captures the effect of the variables of interest and shall be discussed in more detail later on. Again, α_i are individual-specific effects that are not confined to one electoral arena, but determine an individual's innate ability to report voting in *any* electoral arena. This is the unobserved heterogeneity.

A key motivation for using panel data is to prevent bias as a consequence of unobserved or unspecified time-constant factors that may act as confounders. As discussed in Chapter 3, fixed effect models allow for a consistent estimation of the effects of interest, even if unobserved time-constant factors are arbitrarily related to the variables of interest. Unobserved heterogeneity is particularly challenging for limited dependent variables. For binary dependent variables in a panel data structure, such as participating in an election or not, I propose to use Chamberlain's (1980) conditional logit model (PCL). After conditioning on a sufficient statistic s_i, the number of successes for each respondent over all panel time periods,

the conditional maximum likelihood function is free of unobserved heterogeneity (see e.g., Allison, 2009; Lee, 2015). The resulting conditional choice probabilities have the structure $P(y_{it} = 1|s_i) = f(\omega_{it})$, from which the α_is are successfully eliminated. The researcher is left only to specify the factors that vary between the elections ω_{it}.

Parameterizing the election-specific factors

I specify three models that capture the competing theoretical arguments about the causal relationship between EU preferences and electoral participation: Second-order, legitimacy and policy alienation. Importantly, two kinds of factors that vary between the elections need to be distinguished. First, a variable can vary across elections, while its coefficient remains constant. Second, the coefficient of a variable (no matter whether time-varying or -constant) can vary across elections.

First of all, I include election-specific intercepts α_t that captures differences in the overall level of electoral participation between the elections. Second-order theory suggests that the key variable that influences turnout is the importance that respondents assign to elections in a particular arena. Elections in the first-order arena are seen as most important, and second-order elections as less important. Obviously, measures of the electoral importance are specific to the electoral arena. However, as the change of electoral arenas is synonymous with change in time (see Chapter 3), electoral importance can be treated as a time-varying variable. Conceptually, the effect of a unit change in electoral importance should be constant across elections. Accordingly, I model the effect parameter β as constant across elections.

SOE Model: $P(y_{it} = 1|s_i) = f(\alpha_t + \beta \; importance_{it})$

Voter EU and Left-Right preferences are time-varying variables. Both the legitimacy and policy alienation argument however suggest that the effects of policy preferences vary across the elections, as EU preferences become more relevant in EP elections. As neither the legitimacy and policy alienation argument contests that electoral importances matter, their suggested specifications are represented as extensions of the SOE Model specification. The Legitimacy model argues for a direct effect of EU preferences on the likelihood to turn out:

Legitimacy model: $P(y_{it}) = f(\alpha_t + \beta \; importance_{it} + \gamma_{1t} EU_{it})$

Note that γ_{1t} varies over t, so for each election one coefficient is estimated.

The policy alienation model implies that the election-specific effect of EU preferences varies conditional on respondents' Left-Right preferences. This conditionality is incorporated by including an interaction effect between EU and Left-Right preferences. Moreover, the interaction is said to not be linear but U curve shaped. To allow for non-linearity in the conditional relationship, I include an additional interaction between EU and the squared Left-Right variable, and all constituent terms (Brambor et al., 2006; Berry et al., 2012). As the effects of policy prefer-

ences are election-specific, the policy alienation contains 12 more additional effect parameters than the legitimacy model.

Policy alienation model: $P(y_{it}) = f(\alpha_t + \beta \ importance_{it} + \gamma_{1_t}EU_i + \gamma_{2_t}LR + \gamma_{3_t}LR^2 + \gamma_{4_t}EU \times LR + \gamma_{5_t}EU \times LR^2)$

The three panel conditional logit models are estimated by maximizing conditional log-likelihood w.r.t. the α's, β's and γ's.[6]

Theoretical predictions

Second-order theory posits that turnout behavior across electoral arenas is primarily explained by the relative importances of elections in these arenas, and independent of issues that are specific to the arena. For the data to be consistent with the second-order argument, the second-order model should fit the data best. I compare model fit with Likelihood ratio tests, which in this case should be insignificant when comparing the legitimacy policy and the alienation model to the second-order model.

I regard the legitimacy argument as supported by the data if three conditions are fulfilled. First of all, the legitimacy model should fit the data significantly better than the second-order model and the alienation model. Secondly, I expect the effect of EU preferences to be significantly larger in the European arena than in the Federal arena and the State election. Thirdly, the magnitude of the effect of EU preferences should be independent of Left-Right preferences.

The policy alienation model implies that the magnitude of the effect of EU preferences systematically varies conditionally on voter Left-Right preferences. A first condition for this to be the case is that the policy alienation model fits the data significantly better than the legitimacy argument. Additionally, the shape of the effect conditional on voter Left-Right preferences needs to be in line with the constellation of parties in the European policy space.

4.4 Operationalization

Dependent variable Election-specific electoral participation is measured as described in Chapter 3. In the panel conditional logit model, the log-likelihood contribution of uniform turnout sequences, that is of respondents who always or never reported to have turned out, is zero. This follows from the logic of the conditional likelihood approach that uses each respondent as its own control group. Since only within-unit variation identifies the effects of interest, respondents with uniform choice sequences are effectively dropped from the analysis (see Chapter 3 for a detailed discussion). The effective sample size is therefore reduced consider-

---

[6]I maximize conditional log-Likelihood using R's *optim()* function. Monte Carlo simulations show that the estimator converges reliably and fairly quickly on the true parameters.

Voting sequence	000	001	010	011	100	101	110	111
Frequency	54	15	39	42	18	9	416	1837

Table 4.1: Dependent variable: Frequency of individual-level turnout sequences in State, Federal and EP election. 1 indicates reported participation, 0 reported abstention.

ably, since it is confined to the voters with variation on the dependent variables across elections. Table 4.1 shows the frequency of turnout sequences across the State, Federal and EP elections for respondents with no missing information on the independent variables employed in this analysis. 1 signifies that the respondent reported turnout, and 0 otherwise. There are 539 respondents that changed their voting behavior between the three elections, and 1891 with uniform turnout sequences (000, 111). Those reporting to have turned out in the state and the federal, but not the EP election (110) are by far the most common among the non-uniform turnout sequences.

Policy preferences   Policy preferences are operationalized as described in Chapter 3. To facilitate interpretation of the coefficients, I center both EU and LR preferences on the midpoint on the 0-10 scale, 5.

Electoral importance   The MEDW Bavaria panel study repeatedly asked respondents to indicate what importance they assign to State, Federal and EP elections, using a 0 to 10 scale.[7] All three items were part of the survey in all pre-electoral waves, i.e., three times. Variation of electoral importances over the panel waves is minimal. Rather than variation in time, variation between electoral arenas is the defining characteristic of this measure. I therefore chose to average importance evaluations for each respondent and each kind of election over the three waves to obtain a robust measure of the structure of respondent's importance evaluations that is less prone to measurement error. This also has the benefit of reducing missing information considerably - only three respondents that reported to have changed their voting behavior between the arenas had to be dropped due to missing information on this variable. To facilitate interpretation of the coefficient, I finally recode electoral importances as the difference in importance to the importance of the Federal election. This is of no further consequence since only variation in importance across elections is used by the PCL model anyways. Comparing the means, I find that State elections are seen as only slightly less important than Federal elections (0.2). Most respondents clearly identify EP elec-

---

[7]The wording implied that respondent should rate the importance not of a specific individual election, but elections in a specific electoral arena in general: "Please rate the importance to you personally of these three elections: State elections, Federal elections, European elections." The three 0-10 scales ranging from *Not important at all* to *extremely important* were arranged in an item battery.

tions as least important: On average, the difference in importance between Federal election and EP election is 3.1.

Time-varying control variables   While all individual-specific time-constant factors are "controlled" for in the panel conditional logit, there may be time-varying individualistic factors that explain changes in voting behavior across the elections. If these are also causally related to policy preferences, these factors need to be included in the model in order to identify the parameters of interest. To keep my model parsimonious, I focus only on two main factors that have been identified in the relevant literature so far. Second-order theory argues that government popularity affects changes in voting behavior in EP elections. The original argument operates at the party-level (Reif and Schmitt, 1980), and is therefore captured by the election-specific intercepts. However, individuals may without doubt vary in their perception of government performance. Individual-level studies therefore traditionally have included measures of respondents government evaluations to incorporate the individualistic protest motive into their models (e.g., Hobolt et al., 2009; Schmitt et al., 2008). Equally, I include a time-varying measure of satisfaction with the Federal government as a control variable. Respondents were asked two times how satisfied they were with the performance of the "Federal government in Berlin", once before the Federal and once before the EP election. Respondents used a 4-point scale ranging from "Not satisfied at all" to "Very satisfied" to indicate their assessment. Due to the short time-span between State and Federal election, I use the first measurement for the State election as well. Economic factors have been identified as an additional motive for changing voting behavior in EP elections (Braun and Tausendpfund, 2014; Garry and Tilley, 2015; Hobolt and de Vries, 2016; Schakel, 2015). Insofar as these might be also causally related to policy preferences, they constitute potential confounders that should be controlled for.[8] I do so by including a measure that captures changes in the perceived economic situation of the respondents between the elections. Respondents were asked two times if they thought they were financially better off (1), about the same as a year ago (2), or worse off (3) than last year.

4.5   Bivariate analysis

Are policy preferences associated with turnout in EP elections? I conduct a first bivariate investigation of the relationship on the 539 respondents that changed their voting behavior across the three election, and therefore constitute the effective sample. Figure 4.1 plots the proportion of respondents that report to have participated in the EP election, given the respondents' EU preferences as reported

---

[8]For the vivid debate about the causal relationship between economic perception and voting, to which this contribution shall remain largely agnostic, see e.g. Lewis-Beck et al. (2008); Evans and Andersen (2006); Johnston et al. (2005).
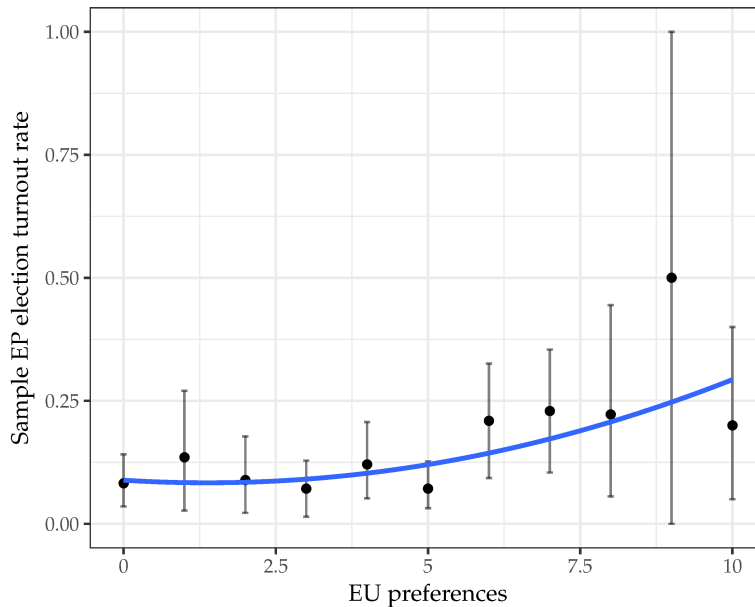
Figure 4.1: Proportion of voters per EU integration score, with 95% confidence intervals. Second-order polynomial linear fit.

at the time of the EP elections. Additionally, I plot the line of best second-order polynomial linear fit.

The proportion of voters among the respondents with low scores (0-4) on the EU integration dimension is notably smaller than among respondents in favor of more EU integration (6-10). While the percentage of voters ranges between 7-13% in the former group, in the latter around 20% are voters. Of course, the statistical uncertainty indicated by the 95% confidence intervals is substantial, since there is only a small number of observations for each unit step on the EU integration dimension. However, I think the overall pattern clearly suggests that there is a relation between EU preferences and turnout. This interpretation is also supported by the second-order polynomial linear fit which slopes upwards. Moreover by a comparison of means: the mean EU integration score among voters is one whole point higher (95% CI: [.27, 1.75]) than the mean EU integration score among abstainers.

While this first investigation suggests that there likely is a bivariate association between EU preferences and EP turnout, one clearly cannot determine from this whether this relationship is causal, or merely due to other confounding factors. Before I estimate the panel fixed effect model that will allow to credibly rule out confounding, I want to first test if there is also an indication in the raw data that the association between EU preferences and turnout is contingent upon Left-Right preferences as predicted by the policy alienation argument. I therefore further subset the data according to the Left-Right preferences. Since there are too few

67
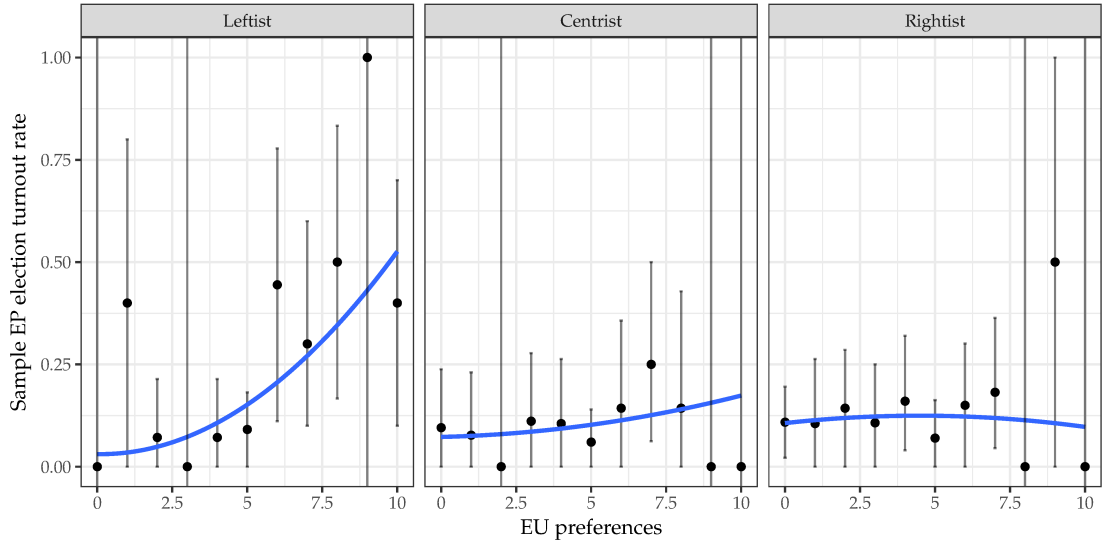
Figure 4.2: Proportion of voters per EU integration score and Left-Right preference grouping, with 95% confidence intervals. Second-order polynomial linear fit.

observations for testing each combination of EU integration and Left-Right score, I group the respondents in only three groups on Left-Right. Of course the grouping will be quite arbitrary and crude, but it will nevertheless produce valuable insights. I group respondent into leftist (Left-Right score 0-4), centrist (5) and rightist (6-10) voters. Figure 4.2 again plots the proportion of voters given their EU preferences, but now separately for each of the Left-Right groups. These are depicted by the horizontally arranged panels.

Although statistical uncertainty about the percentages is huge due to the small number of observations, one can clearly see that the observed patterns are remarkably diverse across the three Left-Right groups. For the group of Leftist voters, more pro-EU preferences seem to be strongly associated with higher participation rates, for centrist voters the association is much smaller, but still notable. There seems to be no association among rightist voters.

Hasty conclusions from this cut at the raw data are clearly not in order. However, I think that the investigation has provided some tentative insights into the relationship between EU preferences and turnout in EP elections. The bivariate investigation suggests that voters with more pro-EU integration self-placements are more likely to turn out. Moreover, the trivariate investigation reveals clear signs that the relationship between EU preferences and turnout is conditional on voters' Left-Right preferences. More pro-EU integration preferences seem to be particularly strongly related to turnout among Leftist voters, and no association can be established for rightist voters. This trivariate relationship seems to be largely consistent with the policy alienation explanation, as, given the unbalanced constellation of political alternatives in the political space, leftist voters should feel

less alienated, and rightist more alienated, if they hold more pro-EU preferences.

4.6 Results of the panel fixed effects analysis

I estimate three main models that are specified as outlined in Section 4.3. Additionally, I include the economic situation and government satisfaction as time-varying control variables for all models. Table 4.2 displays coefficient estimates and standard errors. The second-order model explains election-specific turnout as a function of the importances that voters ascribe to elections in the electoral arena and time-varying controls. As expected, I find that electoral importance has a statistically significant positive effect on the likelihood to turn out.[9] Taking into account that the variable is coded relative to the importance of the Federal election, which has the value of zero, each unit on the 0-10 scale another election is seen as less important leads to a decrease of .35 (95% CI: [.24;.46]) in the linear predictor function for turnout. As the mean importance of EP elections is roughly 3 points lower, this translates in an average effect on the linear predictor of -1.05 [-.7; -1.39]. As PCL gets rid of the individual-specific constant factors, putting this in real terms, i.e., predicted probabilities or average marginal effects, is not possible. However, in order to help grasp the magnitude of the effect, one can think of the effect on a synthetic voter with certain characteristics. Imagine a voter who will turn out with a probability of .8 in any election, which corresponds to her individual intercept being 2.2 in linear predictor terms.[10] In an election that is seen as three points less important, her probability to turn out is predicted to be around 20 percentage points lower, at .58 [.5; .66]. As this constitutes a quite substantial drop in the probability to participate, I assess the magnitude of the effect of relative importances on turnout as substantially meaningful.

---

[9]The effect parameter of importance is modelled as homogeneous over the three elections. I tested this assumption in an auxiliary model (not reported) where I allowed for election-specific parameters. I found only minimal variation between the election-specific importance parameter, a finding which lends supports to the parameter homogeneity assumption.

[10]Furthermore, one would have to also assume that the unobserved individual-specific intercepts, the unobserved heterogeneity, to be unrelated to all included variables.

	SOE	Legitimacy	Legitimacy+	Alienation
α_{state}	0.75*	0.68*	0.72*	0.74*
	(0.19)	(0.20)	(0.23)	(0.23)
α_{ep}	-1.36^*	-1.35^*	-1.21^*	-1.17^*
	(0.18)	(0.19)	(0.22)	(0.22)
Importance	0.35*	0.34*	0.35*	0.35*
	(0.06)	(0.06)	(0.06)	(0.06)
Economy	0.11	0.09	0.08	0.13
	(0.18)	(0.18)	(0.18)	(0.19)
Gov.satisf.	-0.02	-0.01	0.03	0.02
	(0.19)	(0.19)	(0.19)	(0.20)
EU_{state}		0.08	0.08	0.02
		(0.07)	(0.07)	(0.08)
EU_{fed}		-0.00	0.00	0.03
		(0.08)	(0.08)	(0.10)
EU_{ep}		0.13*	0.14*	0.19*
		(0.07)	(0.07)	(0.08)
LR_{state}			-0.14	-0.19
			(0.10)	(0.11)
LR_{fed}			-0.18	-0.24^*
			(0.11)	(0.12)
LR_{ep}			-0.07	-0.15
			(0.09)	(0.11)
LR^2_{state}			0.06*	0.07*
			(0.03)	(0.03)
LR^2_{fed}			0.05	0.05
			(0.03)	(0.04)
LR^2_{ep}			0.02	-0.01
			(0.03)	(0.04)
$\mathrm{EU} \times \mathrm{LR}_{state}$				-0.02
				(0.03)
$\mathrm{EU} \times \mathrm{LR}_{fed}$				-0.04
				(0.03)
$\mathrm{EU} \times \mathrm{LR}_{ep}$				-0.10^*
				(0.03)
$\mathrm{EU} \times \mathrm{LR}^2{}_{state}$				0.01
				(0.01)
$\mathrm{EU} \times \mathrm{LR}^2{}_{fed}$				-0.01
				(0.01)
$\mathrm{EU} \times \mathrm{LR}^2{}_{ep}$				-0.01
				(0.01)
N	539	539	539	539
LogLikelihood	-261.2	-258.5	-254.9	-246.1
p-value LR test		0.14		< 0.001

Table 4.2: Panel conditional logit model estimates with standard errors in brackets. * indicate $p < 0.05$. Likelihood ratio tests are evaluated relative to the model specification in the column to the left.

The legitimacy model additionally includes time-varying EU preferences with election-specific effects. As there are three elections (subscripts *state*, *fed*, *ep*), three additional parameters were estimated. As the second column of Table 4.2 shows, the effect of EU preferences in the State election is estimated as positive, but indistinguishable from zero. For the Federal election, the effect is estimated as close to zero. EU preferences only have a positive effect on turnout in the EP election that is confidently distinguishable from zero. The magnitude of the effect can best be gauged by setting it into relation to the effect of electoral importance: If we put some trust into the point estimate, the effect of EU preferences is roughly one third of the effect of importance, meaning that being three points more in favor of further EU integration (11pt scale) can offset about a one point decrease in importance (also 11pt scale). Since the average drop in importance between the Federal and EP election is about 3, a voter would need to move about 8 points on the EU integration scale. So I would suggest that the estimated magnitude of the effect of EU preferences is not meaningless, but also not very substantial in the grand scheme of things.

This cursory reading of the estimated parameters indicates that my findings of the Legitimacy model may be consistent with theoretical expectations. EU preferences do seem to only play a role in EP, and not in domestic elections. However, my hypothesis are not about the coefficient per se, but whether the effect of EU preferences is larger in EP elections than in domestic elections. Since the difference between a significant and an unsignificant coefficient is not automatically significant itself, I investigate the first difference. I take a large number of random draws from the coefficient sampling distribution and calculate the difference between the EP coefficient and the State and Federal coefficient for each draw. Densities are plotted in Figure 5.4. P-values are approximated by the proportion of coefficient draws for which a negative difference was obtained.

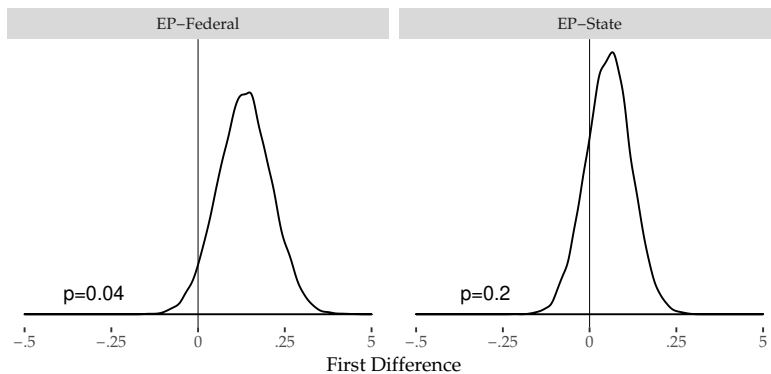The difference in the EU preference coefficients for the EP and Federal elections

Figure 4.3: Legitimacy model: Density of first difference in election-specific EU preference coefficients

is statistically larger than zero, which indicates that EU preferences play a larger role in EP elections than in Federal elections. For the comparison of effect between EP and State elections, one cannot confidently exclude the possibility that the true coefficients are the same. The ambiguity of the findings from the Legitimacy model are further supported by a comparison of the model fit: The Loglikelihood ratio test indicates that the Legitimacy model does not fit the data significantly better than the second-order model. Overall, I therefore conclude that the Legitimacy model, while indicating that EU preferences may play a role in explaining EP-specific turnout, provides ambiguous results that do not satisfy all observable implications drawn from the theoretical argument.[11]

Before proceeding to the policy alienation model, I run a modified Legitimacy model (*Legitimacy+*) which serves as the new base model against which the policy alienation model can be evaluated. To the Legitimacy model specification, I add Left-Right preferences and its square, both of which also enter the specification of the policy alienation model. Accordingly, the Legitimacy+ model differs from the policy alienation model only insofar as it does not include interaction effects between EU and Left-Right preferences. The policy alienation model then tests whether the effect of EU preferences is conditional on Left-Right preferences by including interaction effects. Although interaction effects are cumbersome to interpret from a regression table, it may give us some first indications. First of all, the previously observed patterns remain very robust: Parameter estimates for electoral importance and the direct effect of EU preferences remain substantively the same. Secondly, the interaction effect between EU and Left-Right differences is negative, but is only statistically significant and substantial for the EP election. Since both policy preference measures were centered on 5, moving one unit to the right from the center is estimated to modify the direct EU preference coefficient by about -.1, and moving one unit in Left-Right to the left will increase it by .1. This indicates that the effect of EU preferences is larger for leftist voters than for rightist voters. The interaction effect between EU and squared Left-Right preferences are all inconclusive, indicating that the interaction is rather linear. Finally, the policy alienation model fits the data substantially better than the Legitimacy+ model, and indeed all other models that were estimated, as indicated by the significant Likelihood Ratio test.[12]

I test whether the direct effect of EU preferences is larger in the EP election than in the domestic election.Due to the interaction effect, the interpretation is however different, as the direct effects now captures the effect of EU preferences for a voter who is located at the midpoint of the scale (policy preferences were recentered

---

[11]The Legitimacy model tests for a linear effect of EU preferences on turnout. Since the assumption of effect linearity may not hold, I also tested for non-linearities by including second- and third-order polynomials. I did not find any indications for non-linearities.

[12]I tested for potential effect non-linearity by including also an interaction between squared EU integration scores and squared Left-Right scores, and included all constituent terms. I did not find any indications for non-linearities in the conditional effect of EU preferences on turnout.
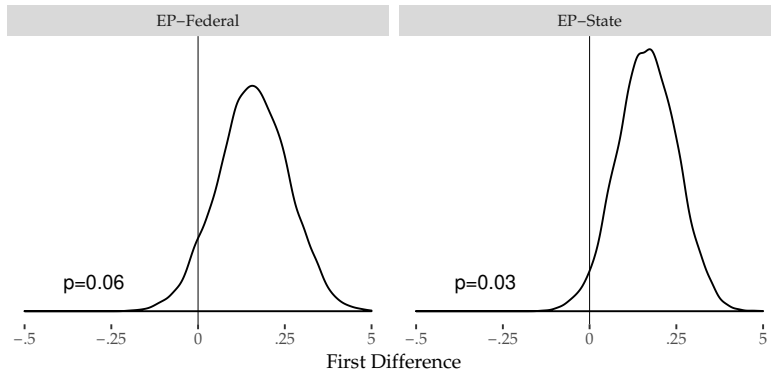
Figure 4.4: Policy alienation model: Density of first difference in election-specific EU preference coefficients for Left-Right centrist voter

on 5, see previous section). Figure 4.4 shows that for the alienation model, the difference in coefficients is barely not significant for the contrast EP-Federal, but significant for the contrast EP-State. To investigate the conditionality on Left-Right preferences more closely, I calculate the marginal effect of EU preferences on the linear predictor function in the EP election, given Left-Right preferences. I do so by taking random draws from the coefficient sampling distributions, and calculating the linear effect of a one unit change in EU preferences, given different values for Left-Right preferences for each draw. Figure 4.5 presents the findings.

For a Left-Right centrist voter (5), becoming one unit more in favor of further EU integration has a significantly positive effect on turnout. The more a voter locates himself to the left of the center, the stronger the effect becomes. For a voter located two units to the left on the Left-Right dimension, the effect is almost twice as high. If a voter is located to the right of the center, the marginal effect of EU preferences becomes indistinguishable from zero. It may even become negative for right-wing voters, however there is substantial uncertainty about this.

I count this as strong evidence that the effect of EU preferences on turnout in EP elections varies systematically conditional on voters Left-Right preferences. The direction of the conditionality is consistent with the predictions derived from the policy alienation argument. However, a final test has to be conducted. The investigation into the results from the policy alienation model has so far only looked at turnout in EP elections. This is not sufficient, since I want to know whether the effect is stronger in EP elections than in other elections. To capture the difference in election-specific effects, I modify the way in which the above presented marginal linear effect was calculated. For each coefficient draw from the sampling distribution, I additionally calculate the marginal effect of EU preferences in the State and Federal election on turnout in these elections, and then subtract these from the ones obtained for the EP election. Thereby, I obtain the first difference of the marginal effect between turnout in the EP election and the Federal election, and between turnout in the EP election and the Federal election.

Figure 4.5: Policy alienation model: Point estimate and 95% confidence interval of marginal effect of EU preferences on the linear predictor function for turnout in EP election.

Figure 4.6: Policy alienation model: Point estimate and 95% confidence intervals of first difference in marginal effects of EU preferences on the linear predictor function for turnout.
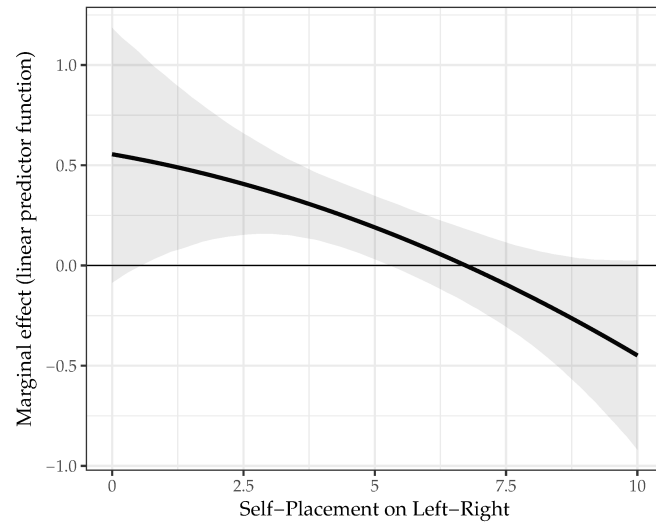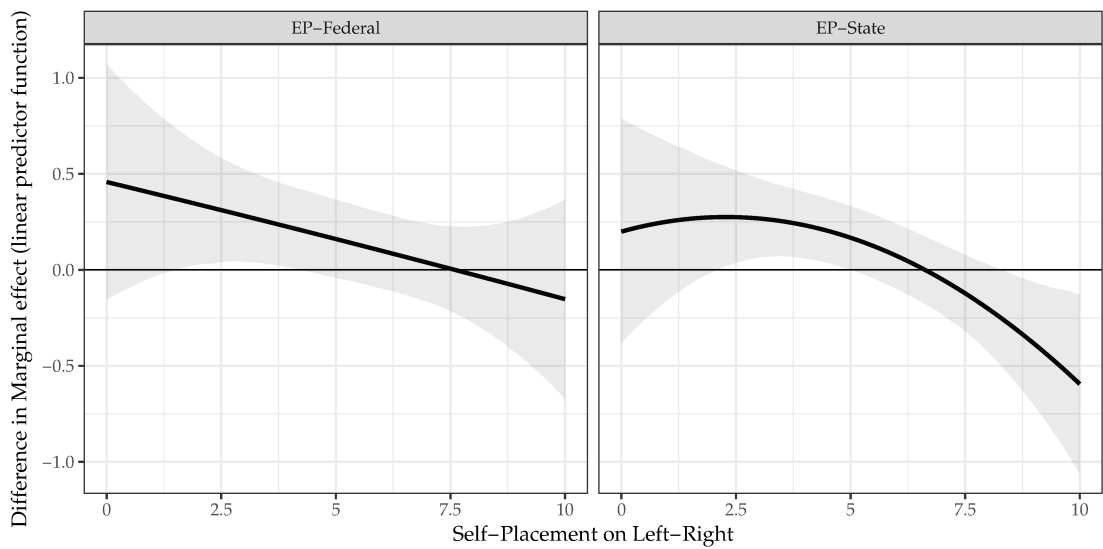
The results are plotted in Figure 4.6. The first difference between the EP and Federal election is depicted in the left panel, the one between EP and State election in the right panel. The substantive patterns that were observed for EP election turnout remain intact. The differencing across elections primarily only introduces additional uncertainty. EU preferences have a larger positive effect in the EP election than in Federal election for Leftist voters. If the contrast is changed to the comparison between the EP election and the State election, I also find that more pro-EU preferences have a larger negative effect in EP elections for right-wing voters.

I conclude that my findings are only consistent with the policy alienation explanation. I do not find that the effect of EU preferences is constant across the elections, as predicted by the second-order model's transfer hypothesis. However, I also do not find that EU preferences have a direct effect on turnout in the EP election, as implied by the legitimacy argument. The increased effect of EU preferences in EP elections seems to be contingent upon voters' Left-Right preferences as predicted by the policy alienation explanation.

4.7 Conclusion

EP election researchers stand divided on whether the views voters hold on Europe-related issues contribute to low turnout rates in EP elections. I test three competing theoretical arguments about the relationship between EU integration preferences and electoral participation. Second-order theory suggests that EU preferences have the same impact on EP election turnout as on first-order election turnout. The legitimacy argument stipulates that EU preferences have a larger impact on the likelihood to participate in EP elections than in other elections, and that the effect is direct and unconditional on Left-Right preferences. A third argument, derived from a multidimensional spatial model of voting, predicts an indirect effect via the route of policy alienation. If European preferences play a larger role in European elections than in national elections, particular ideological subsets of the electorate should be less likely to turn out as they experience higher policy alienation.

I develop a research design to test the competing predictions. My research design focuses on the internal validity of inferences drawn from observational individual-level panel data of Bavarian voters in the election years 2013/2014. I address major inferential problems that plague empirical research on individual-level turnout behavior: retrospective bias, overreporting bias and unobserved heterogeneity. I find that the employed panel conditional binary logit is well suited for the analysis of the determinants of turnout across multiple electoral arenas. While the method has its caveats, it is an attractive option to control for unobserved heterogeneity because of its low data requirements and computational ease. Although my results may not qualify as causal estimates of the relation between EU preferences

and electoral participation, I would argue that they constitute the most credible estimates in terms of internal validity that have been presented so far. I find clear empirical evidence that voter preferences on EU integration have a substantial effect on the likelihood to turn out, but *only* in EP elections. Moreover, the effect appears to be contingent upon the Left-Right preferences of voters. Only among Leftist voters are pro-EU integration views associated with substantially higher turnout, and more anti-EU views with lower turnout. Rightist voters with more pro-integration preferences are not more likely, or maybe even less likely to turn out than rightist voters with anti-EU integration preferences.

The finding that there is an interaction effect between Left-Right and EU integration preferences is remarkable in two ways. First, the finding is substantively interesting, as it gives a novel insight into how preferences structures are related to political outcomes in EP elections. This study is (to my knowledge) the first study that has tested for and discovered such an effect. Secondly, the finding is not consistent with traditional theoretical predictions about the relationship between EU integration preferences and EP election turnout. The second-order elections model predicts that policy preferences have no specific effect on EP election turnout, and the legitimacy argument predicts a direct effect of EU integration preferences that is independent of a voter's Left-Right preference. The finding is however fully consistent with the policy alienation argument derived from a multidimensional model of spatial competition: The effect of EU preferences on turnout seems to be contingent on the relative constellation of voters and parties in the two-dimensional policy space. As policy priorities change between electoral arenas - European issues becoming more important in EP elections than in national elections - particular groups of voters feel more alienated in EP elections than in national elections. In the Bavarian party system under investigation, these are anti-EU center-leftist voters, and pro-EU rightist voters.

What about alternative explanations for my findings? From a statistical perspective, not many alternative explanations for the empirical findings come to mind as my research design has ruled out that the estimated effect is confounded by uncontrolled-for voter characteristics. However, an interesting one follows from a methodological limitation of this study: Although the study controls for individual-specific likelihoods to overreport, the possibility of selective overreporting in the European arena can not be ruled out. Pro-European voters might simply feel more pressured into *reporting* turnout in EP elections than Euroskeptic voters, even though they are just as likely to turn out in reality. Due to absence of validated voting records or long-term multi-arena panel data it is not possible to test this alternative explanation convincingly at the present date. Moreover, it is not apparent to me how selective overreporting might explain the conditionality on Left-Right preferences. From a theoretical perspective, there might be other, not yet explored theoretical explanations for the findings. Ultimately, this study has only provided an indirect test of the predictions of the policy alienation argument on one subnational electorate. The policy alienation argument however implies

that how EU preferences are related to turnout ultimately depends on the constellation of parties in the two-dimensional European space, and is predicted to vary across EU member states according to the configuration of the respective party systems. Therefore, it is imperative to conduct further studies in other EU member countries to test whether predictions hold in these other contexts. Good testing cases for future studies are countries in which the party system is substantially different, i.e., where the predominant Leftist parties are Euroskeptic, and pro-EU parties are rightist. Following the policy alienation argument, the conditional relationship between EU preferences and turnout should be reversed in such party systems, i.e., more pro-EU preferences having a negative effect on the likelihood to turn out for leftist voters, and a positive effect for rightist voters. This study may serve as a methodological blueprint for these much needed further investigations of the connection between EU-related preferences and EP election turnout.

Chapter 5

Transfer or recalibration? EU integration preferences and vote choice

5.1 Introduction

The second-order elections model and the Europe Matters argument make diverging assumptions about the role of issues in EP elections. Second-order theorists maintain that, even though the electoral arena changes, the issues remain the same as in the first-order arena. Voters do not evaluate the issues that are at stake in the SOE arena in their own right, but simply transfer their issue-based decision calculus from the first-order arena to the second-order arena. In effect, this *transfer hypothesis* would mean that, if the issues at stake in SOE are considerably different from those in the FOE, input legitimization through elections does not work in SOE since voters are unable to base their voting decision on the relevant issues.

The Europe Matters argument is more optimistic about the ability of voters to distinguish the different issues that are at stake in different policy arenas. In its weak form, voters are seen as incorporating Europe-related issues into their decision calculus when they vote in EP elections. I call this the weak form since voting based on European issues is fully consistent with the transfer hypothesis, assuming that European issues also play the same role in FOE. In its strong form, voters are thought to adjust their decision calculus between FOE and EP elections by putting a larger emphasis on Europe-related issues in EP elections than in FOE. The argument is best expressed in terms of a spatial model of EU issue voting, in which voters "recalibrate their issue space" between elections (Marsh and Mikhaylov, 2010, 13).

Logically, the transfer hypothesis and the recalibration hypothesis are mutually exclusive. Either voters use the same issue calculus across different electoral arenas,

or they adjust it to the specific electoral arena. Which of the two hypotheses is true is a fundamentally empirical question. Although this question is well known (e.g., Clark and Rohrschneider, 2009; Marsh and Mikhaylov, 2010), so far no study has formulated it precisely and tested it convincingly. Rather, both the second-order and the Europe Matters literature tend to use the transfer or recalibration hypothesis as an (more or less implicit) assumption in their theory. I think that this research question is too important to be treated as a mere assumption, as it is not a simple working assumption but the theoretical linchpin for both strands of research.

In this chapter, I formalize the competing hypotheses from the perspective of a multidimensional model of spatial competition. In a two-dimensional policy space consisting of a European integration and Left-Right dimension, transfer and recalibration can be precisely expressed as the absence or presence of changes of the relative dimensional weight across electoral arenas. The research question then boils down to estimating the election-specific relative dimensional weight parameters. However, estimation is not straightforward. In a Monte Carlo study, I show that conventional statistical models will likely underestimate the change in weights in the presence of unobserved heterogeneity that is related to the spatial preferences of voters. The solution I propose is a fixed-effect multinomial logit model that is particularly well suited for voter panel survey data.

5.2 The spatial voting model with relative dimensional weights

The transfer and recalibration hypotheses can be expressed precisely in the multi-dimensional model of spatial competition (Davis et al., 1970; Enelow and Hinich, 1984b; Hinich and Munger, 1997). In line with the EU issue voting literature (e.g., de Vries, 2010; de Vries et al., 2011; Hobolt et al., 2009; Hobolt and Wittrock, 2011; Hobolt and Spoon, 2012), I start with the proposition that the relevant policy space for European voters consists of two (latent) policy dimensions: A Left-Right dimension, which effectively summarizes policy preferences related to economic and social considerations, and an EU integration dimension, which summarizes policy preferences related to the extent and/or deepening of European integration (see also Chapter 2).

Dimensional weights in a single election

In the cross-sectional case, i.e., voting in a single election, the utility U of voter i voting for a party j decreases with the sum of squared Euclidean distance between

the party position p and voter ideal point v on all relevant dimension.[1]. This spatial component is normalized by the parameter β, which expresses the impact of spatial considerations, relative to non-spatial considerations c.

$$U_{ij} = c_{ij} + \beta[(p_{eu_j} - v_{eu_i})^2 + (p_{lr_j} - v_{lr_i})^2] \tag{5.1}$$

At this stage, an equal weight is assigned to both dimensions, meaning that increasing the distance by one unit decreases the utility to the same extent for each dimension. This inflexible proposition is lifted by introducing a mixing parameter ω that allows the dimensions to vary in their relative weights. ω is defined on the unit interval, and expresses the relative weight of distance on the EU dimension in the spatial component (see e.g., Kedar, 2005, for a similar approach)

$$U_{ij} = c_{ij} + \beta[\omega(p_{eu_j} - v_{eu_i})^2 + (1 - \omega)(p_{lr_j} - v_{lr_i})^2] \tag{5.2}$$

In the extreme cases of ω, the two-dimensional model collapses into a unidimensional one: If $\omega = 0$, distance on the EU dimension has no effect on utility, and the spatial component consists entirely of distance on the left-right dimension (since $1 - \omega = 1$), and vice versa if $\omega = 1$. Accordingly, ω can be interpreted as a percentage: If $\omega = .5$, distances on the EU and left-right dimension have the same impact on utility, they are weighted 50:50 etc. While this mixing parameter formulation makes the inner workings of the spatial model transparent and allows for a straightforward interpretation of the weights, it is not convenient to estimate in the maximum-likelihood framework.[2] More common in the spatial literature is the formulation of dimension-specific spatial effect parameters that fits the linear predictor function of generalized linear models, such as $\beta_{eu}(p_{eu_j} - v_{eu_i})^2 + \beta_{lr}(p_{lr_j} - v_{lr_i})^2$. The linear specification has more desirable statistical properties, such as allowing for unconstrained likelihood maximization. As can be shown, the linear predictor formulation can be easily derived from the mixing parameter formulation. Multiplying out the spatial component in Equation (5.2), the dimension-specific weight parameters β_{eu} and β_{lr} are simply $\beta \times \omega$ and $\beta \times (1 - \omega)$. In turn, this means that relative dimensional weights can be recovered from estimates of the generalized linear predictor function. The relative weight of a dimension is the spatial coefficient of that dimension, divided by the sum of all spatial coefficients[3], i.e.,

---

[1]Subscripts eu and lr indicate the EU and the Left-Right dimension. Additionally, I make the standard assumption that preferences are additive separable. This assumption is lifted in Chapter 6.

[2]Point estimation of ω is relatively straightforward. The unit interval constraint on ω can be included by applying an inverse logit transformation on ω ($\omega^* = 1/(1 + exp(-\omega))$). This allows an unconstrained maximization of the likelihood function. Estimates for ω are then obtained post-estimation by employing the transformation to the maximum-likelihood estimates for ω^*. However, standard error estimates are unreliable for boundary cases, e.g., if $\omega \to 0$.

[3]Proof: Substituting β_{eu} (β_{lr}) with $\beta \times \omega$ ($\beta \times (1 - \omega)$) and canceling, the true statements $\omega = \omega$ ($1 - \omega = 1 - \omega$) are obtained.

$$\omega = \frac{\beta_{eu}}{\beta_{eu} + \beta_{lr}}; \qquad 1 - \omega = \frac{\beta_{lr}}{\beta_{eu} + \beta_{lr}}. \qquad (5.3)$$

The multiple elections case for panel data

To reiterate, the goal of this investigation is to compare the relative weight of EU integration preferences across multi-level elections. To this avail, I reformulate a spatial model of voting in multiple elections for panel data. I extend the cross-sectional utility function by introducing the subscript t $(1, 2...T)$ that denotes the particular election that a voter participates in. These might be domestic SOE, FOE or EP elections.

$$U_{ijt} = c_{ijt} + \beta_{eu_t}(p_{eu_{jt}} - v_{eu_{it}})^2 + \beta_{lr_t}(p_{lr_{jt}} - v_{lr_{it}})^2 \qquad (5.4)$$

The key feature of the above equation is that the spatial parameters are modeled as election-specific. This is vital because it is to be tested whether these vary between elections. The β_t's now denote the impact of distances on the EU and Left-Right dimension in the specific election t. Furthermore, party positions and voter ideal points are also allowed to vary between elections, as elections take place at different points in time.

Non-spatial factors c_{ijt} now also vary between elections. These can be further divided up into an election-varying component and an election-constant component δ_{ij}. δ_{ij} is the unobserved heterogeneity. This is the voter-specific contribution to the utility of voting for a specific party that is independent of the election - the general inclination of a voter to vote for a party irrespective of the election. Examples might be long-lasting party attachment or identification, psychological or socio-demographic factors, or any unobserved individual-specific factor that is constant across elections. Furthermore, there are election-varying non-spatial factors, i.e., all changes in the utility of voting for a party that vary between elections. This can be parameterized as a vector of party-election-specific constants or valence parameters α_{jt}, which capture the overall ebbs and flows of party fortunes across elections, a matrix of individual-election-specific covariates X with party-specific effect parameters γ_j and error term ϵ_{ijt}.[4] The full utility function then is:

$$U_{ijt} = \alpha_{jt} + Z_{it}\gamma_j + \beta_{eu_t}(p_{eu_{jt}} - v_{eu_{it}})^2 + \beta_{lr_t}(p_{lr_{jt}} - v_{lr_{it}})^2 + \delta_{ij} + \epsilon_{ijt} \qquad (5.5)$$

Observable implications

Assuming for a moment that estimates of the dimension-specific weight parameters, $\hat{\beta}_{eu_t}$ and $\hat{\beta}_{lr_t}$ are miraculously obtained, the validity of the competing transfer

---

[4]Additionally, this may encompass also a matrix of election- and party-specific time-varying covariates if such covariates are to be included.

and recalibration hypotheses can be evaluated. In a first stage, the relative dimensional weights of the EU dimension (ω_t) can be calculated. These are obtained for each election t by $\hat{\omega}_t = \frac{\hat{\beta}_{eu_t}}{\hat{\beta}_{eu_t} + \hat{\beta}_{lr_t}}$.

Now, the main hypotheses can be readily evaluated in the classical null hypothesis testing framework. The transfer hypotheses forms the null hypothesis as it predicts no change in the relative weight of the EU dimension across elections,

$H_0 : \omega_{EP} - \omega_{FOE} = 0$ (transfer hypothesis)

The recalibration hypothesis is the alternative hypothesis. It stipulates that the relative weight of the EU dimension is larger in EP elections, and can therefore be formulated as a directed hypothesis.

$H_a : \omega_{EP} - \omega_{FOE} > 0$ (recalibration hypothesis)

and

$H_a : \omega_{EP} - \omega_{State} > 0$.

5.3 Estimating the dimensional weights

The conventional statistical approach for estimating the dimensional weights (βs) in Equation 5.5 is McFadden's (1974) conditional logit model, a generalized form of the multinomial logit that allows for alternative-specific variables such as policy distance between individuals and party positions. This makes conditional logit the workhorse model for empirical applications of spatial models (e.g., Alvarez and Nagler, 1998; Schofield and Zakharov, 2009; Stoetzer and Zittlau, 2015).[5] As the term conditional logit is extremely confusing in the context of this dissertation due to other similarly named models, I shall refer to McFadden's conditional logit simply as multinomial logit (MNL) in the following.

One possible research design to estimate the weights would be to estimate separate cross-sectional MNL models for each election t. A cross-sectional MNL model can be derived by selecting a particular t in Equation 5.5, and by assuming the error-term ϵ_{ij} to be assumed distributed i.i.d. Type-1 extreme value (McFadden, 1974). Collecting the systematic terms of the utility function in V_{ijt}, the choice probability of voter i for party j is given by $Pr(Y_{ij} = 1) = \frac{e^{V_{ij}}}{\sum_{j=1}^{J} e^{V_{ij}}}$. Theoretically, parameter estimates, obtained by maximizing likelihood, can then be readily compared. However, this is only possible if very strong assumptions are met. In practice, parameter estimates from discrete choice models such as multinomial logit are not comparable across samples, groups within samples or over time (Mood, 2010; Allison, 1999; Wooldridge, 2010, 582-585). Unlike linear models, discrete choice estimates are affected by omitted variables, even if these variables are unrelated

---

[5]MNL relies on the independence of irrelevant alternatives assumption. Multinomial probit is considered a solution (Schofield et al., 1998; Quinn et al., 1999; Dow and Endersby, 2004), but is computationally unwieldy. Practical experience shows that the consequences of IIA violations are often minor, especially in stable party systems (Whitten and Palmer, 1996, 255).

to the variables included in the model. At the root of the problem is unobserved heterogeneity which can bias estimates upwards or downwards to a different degree in different samples. Estimates from multiple cross-sectional models will therefore only be comparable if the model is fully and correctly specified - i.e., the entire voting calculus of voters is captured in the correct functional form - certainly a hopeless endeavor. This comparability problem is ameliorated if the individuals remain the same as is the case in panel data set. However, the fundamental problem only goes away with panel data if the omitted variables are unrelated to the included variables, as I will show in the Monte Carlo study below.

The conditional likelihood approach offers an elegant way to circumvent these problems, which will allow to draw stronger inferences about the variation of relative weights across elections. As has been noted already by Chamberlain (1980), the conditional likelihood approach is not limited to binary logit but also extends to MNL. Compared to the binary case, the panel conditional multinomial logit (PML) has one more dimension (J alternatives), which complicates the model somewhat. However, the logic of conditioning on a sufficient statistic remains the same. As the reader will draw little additional benefit from a full mathematical exposition of the multinomial extension, this shall be avoided there.[6] In the case of multinomial choice, it can be equally shown that the frequency of successes (in my case the number of vote choices for party j), $s_{ij} = \sum_{t=1}^{T} y_{ijt}$, where y_{ijt} is a dummy indicator of vote choice of individual i for party j at time t, is a sufficient statistic for δ_{ij}. Intuitively speaking, sufficiency means that the aggregate choice pattern, i.e., the sum of the dependent variables, is generated solely by the unobserved heterogeneity, while the time-varying factors generate the *timing* of the choices among the alternatives, the specific party choice sequence. Conditioning on the sufficient statistic in effect means analyzing the choices for a given s_{ij}. For example, if $s_{ij} = 1$, meaning voter i voted for party j once, one estimates in which election this vote choice took place, given that it is known that the choice sequence belongs to the set of all possible choice sequences that could have generated exactly one vote for party j. Since timing becomes the choice variable, time-constant characteristics, of which unobserved heterogeneity is a part, are irrelevant and drop out.

The structure of conditional multinomial choice probabilities $P(y_{ij}|s_{ij})$ resemble the choice probabilities of the MNL, only that the denominator is the sum of the exponentiated linear predictor of all sequences that could have produced the observed aggregate voting pattern (Chamberlain, 1980, 71; Lee, 2012). These sequences are the set of all permutations of the dependent variable, which is sometimes referred to as the permutation matrix (Pforr, 2013). An example might come in handy here.[7] Consider a three-wave panel, in which a voter chooses between three parties. This voter chooses party 1 in the first wave, party 2 in the second and party 3 in the third. Since the voter selected each party once, the sufficient

---

[6]A detailed discussion can be found in Pforr (2013), or Lee (2012), Lee (2015).

[7]This following example is taken from Pforr (2013, 45).

statistics are $s_1=1$, $s_2=1$ and $s_3=1$. The voter's permutation matrix then is the set of all potential sequences that are consistent with the sufficient statistics, i.e., $\{(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)\}$.

Computationally, panel conditional multinomial logit (PML) is very similar to standard MNL and estimation in the maximum-likelihood framework is numerically stable and straightforward. The likelihood function is the sum over all realized conditional probabilities. As in the binary case, the likelihood function is free of the δ_{ij}'s, which means that parameters are consistently estimated even if unobserved heterogeneity is related in any way to the included variables.

A Monte Carlo Study of the Panel multinomial estimator

I conduct a Monte Carlo study to illustrate the consequences of unobserved heterogeneity for a cross-sectional analysis of relative dimensional weights and to investigate the statistical properties of the proposed PML. The scenario under investigation consists of 1000 voters that participate in three elections. The choice set consists of three invariant parties in each of the three elections.

In a first step, the true data generating process of party utilities is specified, following the systematic part of equation 5.5, as a function of

1. party- and election-specific valence
2. party- and election-specific policy distances on two dimensions with election-specific spatial effect parameters (the spatial component)
3. one individual- and election-specific covariate with party-specific effect parameters (the control variables) and
4. one individual-specific, election-constant covariate (the unobserved heterogeneity).[8]

The true spatial parameters are set to -1 for the EU integration dimension and -2 for the Left-Right dimension in the first election, to -2 and -1 in the second, and again to -1 and -2 in the third election. The true relative weights are then 1/3 in the first and third election, and 2/3 in the second election.

In a next step, I incorporate the two possible assumptions about the causal relationship between included and unobserved factors. As discussed above, unobserved heterogeneity can either act as a confounding factor by being related to the included variables, or be unrelated. I generate two versions of the data: In the first version, there is no confounding: Policy distances and election-varying covariates are unrelated to the unobserved heterogeneity - they are strictly exogenous.[9] As I laid out in the previous section, we expect to obtain biased estimates even

---

[8]The stochastic part ϵ_{ijt} is accounted for by the random multinomial draw, from which the dependent variable is generated.

[9]Generated by random draws from a normal distribution.
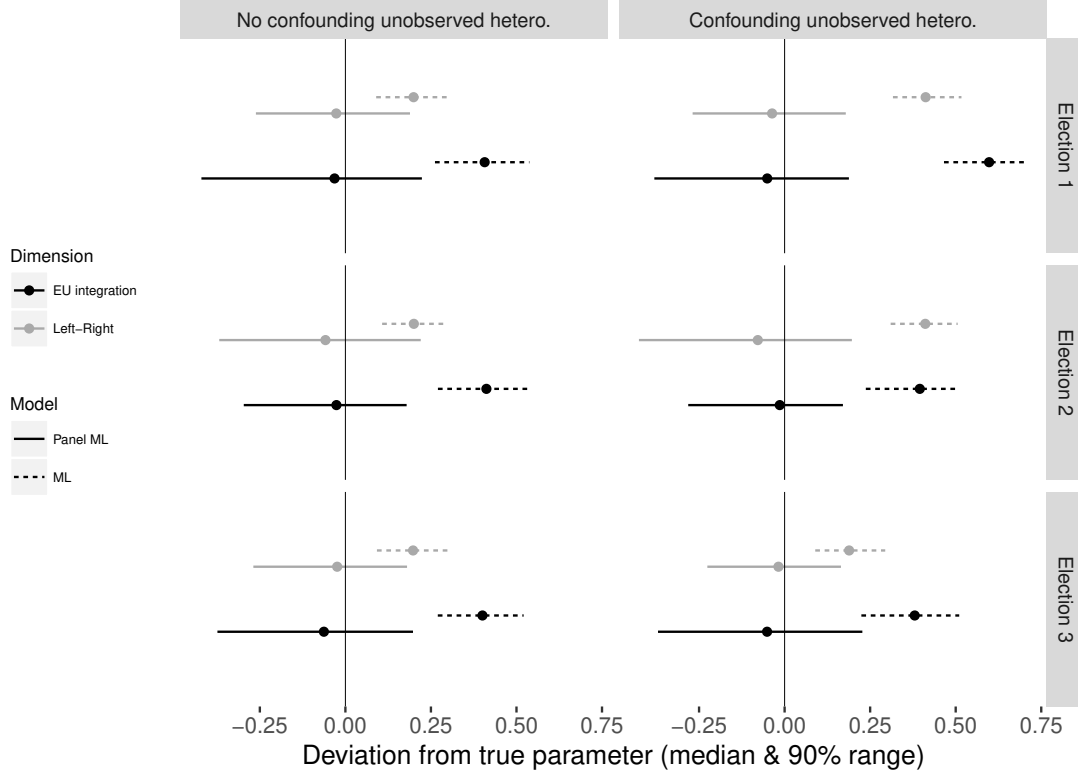
Figure 5.1: Monte Carlo Experiments: Spatial effect parameters

for unconfounding unobserved heterogeneity in non-linear models such as logit. Whether estimates of the relative weight are also biased is not so straightforward, as we shall see. In the second version, policy distance on one dimension is related to the unobserved heterogeneity, and acts as a confounding factor.[10] Here, we expect coefficient and weight estimates to be biased.

In a third step, the calculated systematic utilities (V_{ijt}) are transformed into vote choices. For each respondent in each election, a probability vector containing the party-specific vote choice probabilities is obtained via $Pr(Y_{ijt} = 1) = \frac{e^{V_{ijt}}}{\sum_{j=1}^{J} e^{V_{ijt}}}$. Election-specific vote choices are then generated by a random draw from a multinomial distribution, where the probability vector constitutes the success parameter vector (see Stoetzer and Zittlau, 2015, for a similar approach).

This entire process is repeated 240 times, resulting in 480 datasets with 1000 vote choices in three elections each. For each dataset, I estimate separate cross-sectional multinomial models for each of the three election, and the PML model. Obviously, since the heterogeneity remains unobserved, it is not included in the model specification.

---

[10]Policy distance on this dimension is the sum of a random normal draw and a realization of the unobserved random variable.
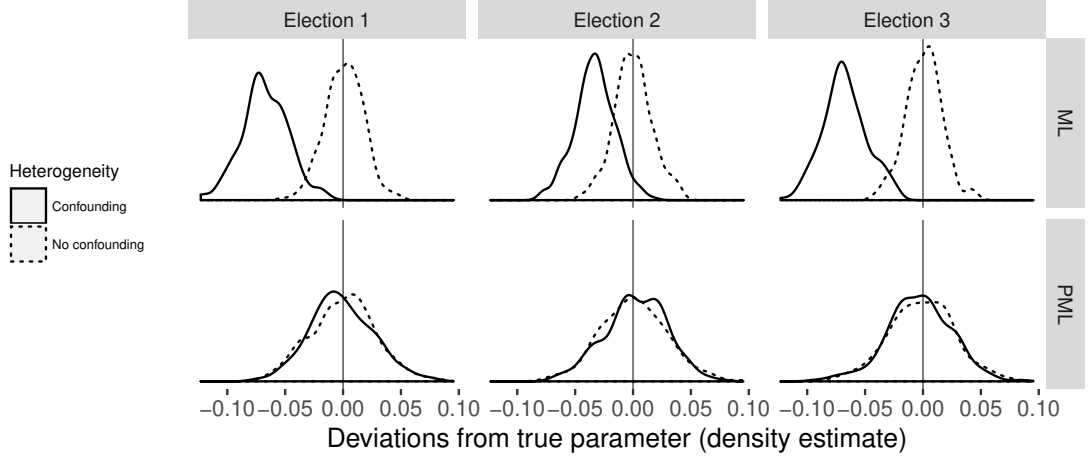
Figure 5.2: Monte Carlo Experiments: Relative weights

Figures 5.1 presents the distribution of the deviations from the true spatial parameters, i.e., the spatial coefficient estimates that I obtained in each run of the Monte Carlo experiment, minus the true parameter value. The distribution of the deviations of the estimates is summarized by its median and 90% range.[11] The findings for the case where unobserved heterogeneity is not a confounder are depicted on the left-hand side, the confounded case on the right-hand side.

First of all, it can be seen that the Panel ML is able to recover the true value of the parameter in all cases. However, the sampling variance of the Panel ML estimator is clearly larger due to the conditioning on the sufficient statistic. As expected, the cross-sectional ML yields biased estimates of the spatial parameters in both cases. It is to be noted that in the not-confounding case, the bias is constant across the elections although the imaginary voters change the relative dimensional weight between the first and second election, and second and third election, whereas the bias is not constant in the confounding case. To investigate this more closely, I calculate the relative dimensional weight obtained in each Monte Carlo experiment. The density function estimate of all weights that approximates the sampling distribution of the weight estimator is depicted in Figure 5.2.

As was hinted at by the constant bias of the spatial parameters, cross-sectional ML recovers the true relative weight in each election if unobserved heterogeneity is not confounding. The reason for this is that spatial estimates for both dimensions are equally affected: The degree to which the effect of distances on one dimension is mis-estimated is proportional to the degree to which the effect of distances on the other dimension is mis-estimated. Therefore the relative weight estimator is unaffected by unobserved heterogeneity that is unrelated to the included variables.

---

[11]MNL intercept term and election-varying coefficient estimates are biased as well, but not reported because they are of minor interest here.

However, this is clearly not the case if unobserved heterogeneity is related to spatial distances, a scenario that is much more likely than the special case of exogenous heterogeneity. In the confounded case, cross-sectional ML yields estimates of the weights that are not comparable across elections. Neither does cross-sectional ML yield unbiased weight estimates, nor is the bias constant across elections. If the "true" relative weight changes between the elections (here from $\frac{1}{3}$ to $\frac{2}{3}$ between Election 1 and 2, and back to $\frac{1}{3}$ in Election 3), so does the extent of the bias. As true changes in the weights can not be separated from changes in the bias, inferences about the first difference of weights between elections are invalidated by unobserved heterogeneity that is related to the included variables. Contrarily, the proposed Panel ML delivers unbiased estimates of the dimensional weights in the presence of unobserved heterogeneity even if it is related to the policy distances. However, this comes at a price: As can be seen by the wider spread of the sampling distribution of the weight estimate, the unbiased Panel ML estimator is less precise than conventional ML estimates.

To recapitulate the main findings so far: Research designs that seek to compare estimates of the relative weight of EU integration issues across domestic and European elections using cross-sectional vote choice data are likely to come up with the wrong answers. The reason is unobserved heterogeneity which prohibits the comparison of estimates across elections. Multi-election panel data alone is not the solution. Even if the unobserved factors are constant within individuals across elections, estimates of relative dimensional weights are not comparable across these elections if unobserved factors are somehow related to the spatial measures. Weight estimates obtained by conventional statistical models are only comparable if the assumption of strict exogeneity holds - in my opinion an assumption that researchers should not be willing to make if it can be avoided. As my Monte Carlo study has shown, more elaborate statistical techniques are required to obtain comparable estimates. The proposed PML can help a great deal to put inferences on a more solid footing by addressing the unobserved heterogeneity problem. Its low data requirements make it an especially attractive option for the short panels that are available. In the next section I discuss the model specification and operationalization of the data.

5.4 Operationalization

Vote choice

I measure vote choice by reported post-election vote choice as outlined in the description of the data in a previous chapter. The AfD did not compete in the State election. This might raise concerns about a potential violation of the IIA assumption upon which the multinomial logit relies. In order to prevent that changes in the the party set due the entry of the AfD dilute the results, I chose to identify counter-factual AfD voters in the State election. I overwrite reported

vote choices only in the State election in instances where respondents indicated a certain hypothetical vote choice had the AfD participated in the election. This procedure affects the dependent variable only in 126 cases of vote choice in the State election, and is of very minor consequence for the results.[12] In the following, I refer to this operationalization of the dependent variable as *Observed Choice*.

The dependent variable contains 2318 missing values. Missing values are primarily the result of respondents not participating in elections. These missings might be problematic as they might introduce selection bias. Most (1300) stem from the EP election, where turnout was substantially lower. This might lead to my results speaking only about those voters who turned out in the EP election which might be substantively different than non-voters. Moreover, the variables of general interest, the spatial preferences of voters on the EU and Left-Right dimensions, have an influence on the decision to participate, as I show in Chapter 4. In order to prevent these detrimental consequences and to improve the robustness of my results, I generate an additional version of the dependent variable, for which I try to infer the counter-factual vote choice had the respondents participated as good as possible. Effectively, this version of dependent variable tries to simulate full turnout in all elections. In a first step, I overwrite missing reported post-election vote choice with reported pre-election vote intentions. In a second step, I make additional use of a survey question to infer hypothetical vote choice in the EP election for the remaining missing votes. All respondents that reported to have abstained in the EP election were asked which party they would have voted for had they participated. They received the same party set and answer options as those who reported to have participated. I overwrite the remaining missing vote choice information in the EP election with the reported hypotheticals. In the following, I refer to this operationalization of the dependent variable as *Full turnout*.

Policy preferences

Voter ideal points and party positions on the Left-Right and EU dimensions are operationalized as described in Chapter 3. Policy distance in election t is simply coded as the squared Euclidean distance between party positions and voter ideal point in election t.

---

[12]Reported vote choices in the State election are replaced with "AfD", if the respondent indicated that he/she would have certainly voted AfD. The survey question reads: "The party 'Alternative for Germany' did not participate in the Bavarian state election. If this party had participated, would you have voted for it?" Respondents could choose between "Certainly would have voted AfD", "Perhaps would have voted AfD", "Certainly would have not voted AfD" and "Don't know". I also estimated the model without this manipulation of the dependent variable, and get the same results.

Election-varying control variables

While PML allows to treat all factors that are constant across elections as unobserved and to be conditioned out, election-varying factors still need to be specified and parameterized in order to identify the spatial estimates. Conceptually, election-varying factors can be separated into variation due to elections taking place at different points in time, and variation due to the elections taking place in different electoral arenas. Empirically, these two dimensions of variation are virtually impossible to separate since the counter-factuals remain unobserved. In the panel data at hand, one does not observe voting in the same arena at a different point in time, or voting in different electoral arenas at the same point in time. Theories of voting behavior can however serve as guidelines to identify election-varying factors. The task is furthermore facilitated by the fact that the timely distance between the election is relatively small. As the Federal election taking place one week, and the EP election nine months after the State election, I think it is plausible to assume that variation due to slow-moving factors such as partisan realignment, social status etc is negligible. This is further supported by the qualitative evaluation of the political context as a period of political stability (see Chapter 3).

The first major part of trying to capture the party utility variation across elections are the intercept terms. As outlined in the previous section, these are specified as election- and party-specific. Conceptually, the intercept captures the utility of casting a vote for a particular party in a particular election that is constant over the respondents. As is generally the case in multinomial models, the constant of one party needs to be set to zero in order to identify the model - the reference or base category. However, this is not sufficient in PML since only changes in the party intercepts across elections are identified. Additionally to choosing a reference party, a reference category in terms of elections needs to be selected. This is done by additionally setting the intercept of all parties for one election to zero. The remaining free intercepts to be estimated then express the change in party utilities between the elections, relative to the changes of the utility of the reference party. Which empirical phenomena are captured by the intercept terms? The intercepts describe the overall electoral fortunes of the parties across elections. To help the readers' intuition - if we were to estimate the model only with the intercepts on the right-hand side, the point estimates would describe the changes in the observed vote shares of the parties between the elections. Empirically, this encompasses variation in institutional context that favors particular parties in particular electoral arenas, or factors located at the party competition level, or changes in party vote shares over time. Put differently, we can say that the intercepts capture all institutional, time and party-level factors that play out equally in the vote choice calculus of all respondents. It is to be noted that the intercepts therefore implicitly capture many aggregate-level explanations put forward by second-order election theorists to describe the variation of party fortunes between first- and second-

order elections - such as swings in the overall party popularities between elections, whether a party is a government or opposition party, or a small party that benefits from the absence of strategic voting incentives in second-order elections, or a conservative/green/socialist/Euroskeptic party (e.g., Hix and Marsh, 2011, 2007; Marsh and Mikhaylov, 2010; Reif and Schmitt, 1980).

While aggregate-level variation is captured by the intercepts, second-order theory also identifies time-varying individualistic factors that explain voting behavior between first- and second-order elections. If these are also causally related to policy preferences, these factors need to be included in the model in order to identify the parameters of interest. To keep my model parsimonious, and since additional control variables in multinomial choice models quickly burn through degrees of freedom, I focus only on two main factors that have been identified in the relevant literature so far. Second-order theory argues that government popularity affects changes in voting behavior in EP elections. The original argument operates at the party-level (Reif and Schmitt, 1980), and is therefore captured by the election-varying intercept. However, individuals may without doubt vary in their perception of government performance. Individual-level studies therefore traditionally have included measures of respondents' government evaluations to incorporate the individualistic protest motive into their models (e.g., Hobolt et al., 2009; Schmitt et al., 2008). Therfore I include an election-varying measure of satisfaction with the Federal government as a control variable. Respondents were asked two times how satisfied they were with the performance of the "Federal government in Berlin", once before the Federal and once before the EP election. Respondents used a 4-point scale ranging from "Not satisfied at all" to "Very satisfied" to indicate their assessment. Due to the short time-span between State and Federal election, I use the first measurement for the State election as well. Economic factors have been identified as an additional motive for changing voting behavior in EP elections (Braun and Tausendpfund, 2014; Garry and Tilley, 2015; Hobolt and de Vries, 2016). Insofar as these might be also causally related to policy preferences, they constitute potential confounders that should be controlled for.[13] I do so by including a measure that captures changes in the perceived economic situation of the respondents between the elections. Respondents were asked two times if they thought they were financially better off (1), about the same as a year ago (2), or worse off (3) than last year.

Model estimation

I estimate a PML model for each of the two versions of the dependent variables, *Observed Choice* and the hypothetical *Full turnout* vote choice. The linear predictor function is specified according to utility function 5.5, with the exception of δ_{ij}

---

[13]For the vivid debate about the causal relationship between economic perception and voting, to which this contribution shall remain largely agnostic, see e.g. Lewis-Beck et al. (2008); Evans and Andersen (2006); Johnston et al. (2005).

that drops out in the PML. Accordingly, I estimate the party- and election-specific intercepts, the election-specific weights of squared distances on the Left-Right and EU dimension and the party-specific effects of election-varying measures of Government satisfaction and Economic situations. For identification purposes, the intercepts of all parties are set to zero for the first time period, and for one party, the CSU, for all time-periods. As described above, the remaining free intercepts express the change in party utilities between the elections, relative to the changes of the utility of the reference party.

5.5 Results

Table 5.1 displays the point estimates and standard errors of the spatial coefficients.[14] Coefficients express the (average) weight of squared distances on the EU and Left-right dimension in the spatial utility function of the respondents in each election. The first column presents the estimates from the *Observed Choice* model, the second from the hypothetical *Full turnout* model that tests for a potential sensitivity of estimates due to selection on electoral participation. The weight of EU distances is estimated at close to zero and statistically insignificant in both the State and Federal election and in both models. This indicates that distances on the EU dimension have no detectable relevance for vote choice in these elections. In these elections, only distances on the Left-Right dimension are estimated to be a determining factor. The coefficients are statistically distinguishable from zero, with the exception of the Federal election coefficient in the *Observed Choice* model.[15] The spatial coefficient estimates for the EP elections seem substantively different. Here, the estimates for both dimensions are statistically significant, indicating that both dimensions seem to matter in the spatial utility calculus. Moreover, the magnitude of the coefficients is larger for both dimensions than in the other elections, indicating that spatial considerations as a whole played a larger role. This is in line with the second-order theory prediction of more sincere voting in EP elections.

In order to enable a clear-cut testing of the observable implications I have formulated, I calculate the relative weight of the EU integration dimension from the estimated spatial coefficients. I take 10,000 draws from the sampling distribution of the coefficients to incorporate the estimation uncertainty about the coefficients.[16] For each election, I calculate the weight parameter for each draw as $\omega_t = \frac{\beta_{eu_t}}{\beta_{eu_t} + \beta_{lr_t}}$. However, as the unit interval constraint on the relative weight that enables an

---

[14]A complete table of estimates can be found in the Appendix to Chapter 5, in section 5.6.

[15]I can only speculate why the Left-Right coefficient is not distinguishable from zero in the Federal election. One possibility would be strategic voting considerations (see e.g., Gschwend et al., 2016).

[16]The asymptotic sampling distribution is multivariate-normal, with coefficient maximum-likelihood estimates as mean vector, and the inverse of the observed Fischer information as the variance-covariance matrix (Pawitan, 2013).

Election	Dimension	Observed Choice		Full Turnout (hyp.)	
		β	s.e.	β	s.e.
State	EU	-0.001	(0.011)	0.004	(0.008)
	LR	-0.029*	(0.014)	-0.029*	(0.011)
Federal	EU	-0.005	(0.01)	-0.011	(0.008)
	LR	-0.016	(0.013)	-0.031*	(0.010)
EP	EU	-0.038*	(0.011)	-0.033*	(0.008)
	LR	-0.041*	(0.013)	-0.045*	(0.011)

Note: Table reports point estimates and standard errors in parenthesis. * : $p < .05$

Table 5.1: Spatial parameter estimates

interpretation of the weight as a percentage was not taken into account during the estimation procedure, coefficient values in some draws need to be transformed. This is the case when a positive spatial coefficient was obtained in a draw. Larger distances having a positive utility contribution is logically inconsistent with the spatial theory of voting, and prevents an interpretation of the relative weights. Therefore I replace all positive coefficient values with close to negative zero values (-10^{-10}). Effectively this means that draws with a positive EU integration coefficient are interpreted as cases with a relative EU weight of close to zero. While this might seem arbitrary for some readers, hopefully only at first sight, there are two main reasons why I am convinced that this is appropriate. Firstly, a similar procedure would have been employed if the constraint had been implemented during the estimation procedure. Secondly, it leads to a more conservative test as positive coefficients are moved into the direction that is unfavorable to the research hypothesis that I test.

The approximated sampling distributions of relative weights that I obtained by the above described procedure are plotted in Figure 5.3. Horizontally, the panels display the results for the Observed Choice and hypothetical Full Turnout model. The election-specific weight estimates in the three elections are arranged vertically. As could already be inferred from the table of parameter estimates, the relative weight of distances on the EU dimension is estimated at close to zero in the State Election in both the Observed Choice and Full Turnout model. In the Federal election, the relative weight estimate seems ambiguous in the Observed Choice model. The density estimate stretches over the whole continuum, is however also tilted towards zero. The ambiguous picture is a results of the non-significant coefficient for left-right distances in the model. The estimate from the Full Turnout
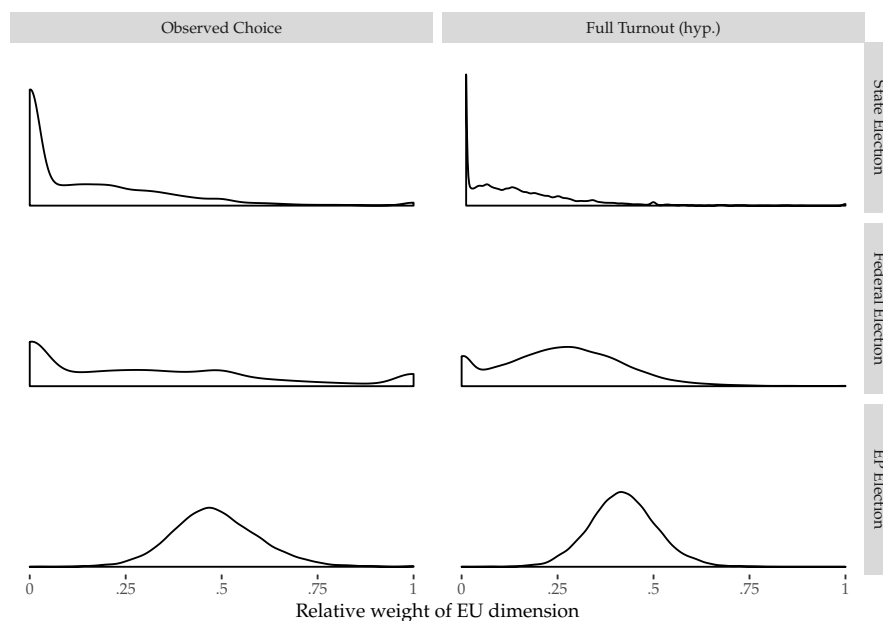
Figure 5.3: Density of Relative weight estimate

specification seems more reliable here. The point estimate of the relative weight of the EU dimension is at around .25, suggesting that voters weighed EU issues at around a quarter, compared to three quarters on Left-Right. However, the spread of the sampling distribution ranges from zero to roughly .5, indicating substantial amounts of uncertainty. My findings are much clearer for the EP election. Here the distribution of the weights peaks close to the midpoint of the continuum, suggesting that voters weigh EU and Left-Right roughly equally. The Full Turnout estimate seems to be a bit more conservative than the Observed Choice estimate. Taking a summary view, my results suggest that the relative weight of the EU dimension seems to indeed vary across the elections. Although there is considerable estimation uncertainty, the pattern of variation seems to be consistent with the recalibration hypothesis. The relative weight seems to be higher in EP elections than in the State and Federal election.

Although the relative weights illustrate the variation of relative weights nicely, a further step needs to be taken to enable a test of the competing hypotheses. To reiterate, the transfer and recalibration hypothesis are not concerned with the relative weights per se, but the change of relative weights between the EP and other domestic elections. While the transfer hypothesis argues that a change of relative weights does not take place, the recalibration hypothesis states that the relative weight of the EU dimension is larger in EP elections than in other elections. Therefore the sampling distribution of the first difference of the relative weights between the EP and State, and EP and Federal election, needs to be investigated. The first differences are simply calculated as the relative weight
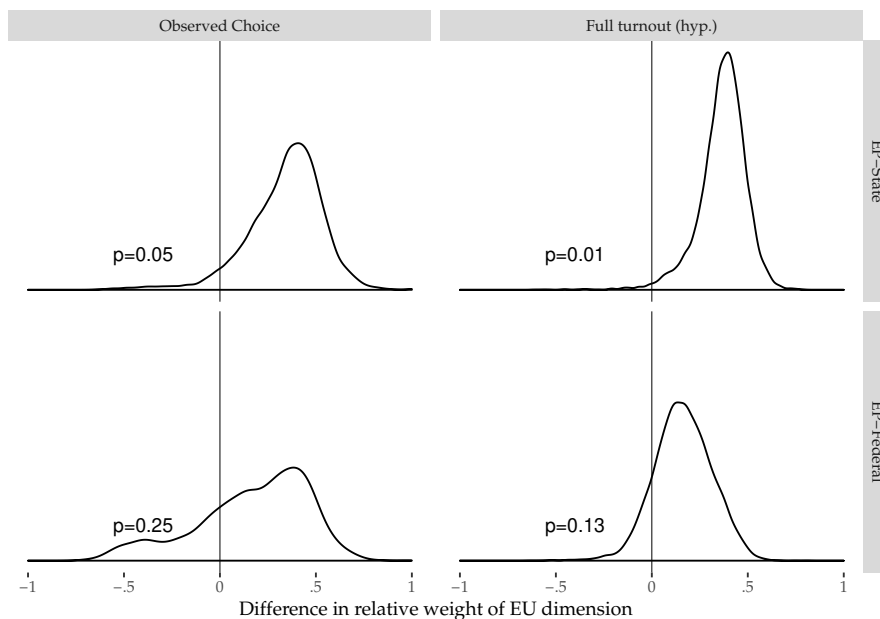
Figure 5.4: Simulated first difference of relative weight and p-values

in the EP election minus the relative weight in the State election, and minus the relative weight in the Federal election for each simulation draw. Note that this procedure is asymptotically equivalent to null hypothesis significance testing (NHST) - the long-run relative frequency of negative differences can be interpreted as the p-value. To calculate the p-value, I simply count the simulation draws in which the first difference is smaller than zero, and divide this count by the overall number of draws (10,000).

Figure 5.4 displays the density of first differences and the simulated p-value. As above, the two estimated models are arranged horizontally, and weight differences between elections vertically. On a first view, one can see that a clear majority of the simulation draws yield positive differences. Positive differences indicate a larger weight in the EP election and are therefore the simulation draws that are consistent with the recalibration hypothesis. However, there is a substantial proportion of the draws with non-positive weight differences, indicated by the density below zero and the p-values. For the Observed Choice model, slightly more than 5% of the simulation draws yield a negative difference between the relative weight in EP and State election. Employing the conventional significance level for NHST, the difference therefore does not qualify as statistically significant. The difference in weights between EP and Federal election are also clearly not significant, as 25 and 13% of the simulations have negative differences. Only the first difference between EP and State in the Full Turnout model is confidently distinguishable from zero.

94

Weighing the evidence

As I have shown above, the transfer hypothesis cannot be consistently rejected. If one would pursue a strict falsificationist standpoint, there is little additional knowledge that can be drawn from this study. As this study fails to reject the null hypothesis consistently, strict falsificationists might simply book this study as evidence for the transfer hypothesis, and move on. However, I don't think that such an approach does justice to the findings. There are good epistemological reasons for a more nuanced interpretation of the findings. I think that in this case, strictly employing the null hypothesis testing framework unfairly stacks the deck in favor of the transfer hypothesis. Any failure to reject the null hypothesis since the p-value falls below an arbitrary threshold, even if due to lacking statistical power, is counted as evidence against the recalibration hypothesis. However, I would argue that the transfer hypothesis is not a null hypothesis that researchers should be happy with. For what it's worth, the transfer hypothesis is not equally plausible as the recalibration hypothesis as it does not built on a convincing theoretical argument. As I laid out in Chapter 2, it was developed rather ad-hoc as a "ceteribus paribus" assumption that enabled second-order theorists to treat policy issues as fixed between elections while they studied the consequences of the change in election importance on voting behavior. It does not logically follow that the "less-at-stake" logic of SOEs necessarily crowd out considerations about which issues are at stake in which election.

Taking the theoretical plausibility of the competing hypotheses into account and embracing the uncertainty that is associated with the findings, a more informative picture arises. The patterns that are observed in the data indicate that the spatial voting calculus of Bavarian voters likely varied between the elections. Indeed, the observed patterns are more consistent with the recalibration hypothesis than the transfer hypothesis. The point estimates indicate a substantial change of the relative weights from close to zero in the State election to roughly a quarter in the Federal election, up to slightly below one half in the EP election. This is in line with the EU issue voting model, which suggests that EU issues play no role in electoral arenas without an electoral connection to EU-level policy making (such as State elections), a diminished role on the national level where an indirect electoral connection exists, and a heightened role in EP elections.

Figure 5.5 illustrates this by plotting the estimated shape of the spatial utility indifference contours.[17] In the State and Federal election, the utility indifference ellipsoid is "compressed" along the left-right dimension, meaning that voters are very critical to deviations along this dimension, and fairly lenient to deviations along the EU dimension. In the EP election, utility indifference contours become almost circular, indicating that both dimensions are weighted roughly equally.

---

[17]Utility indifference contours are calculated using the negative of obtained spatial weight point estimates as the diagonal entries of the A matrix (Davis et al., 1970; Stoetzer and Zittlau, 2015). Additionally, I use weighted Euclidean distance as the spatial metric here.
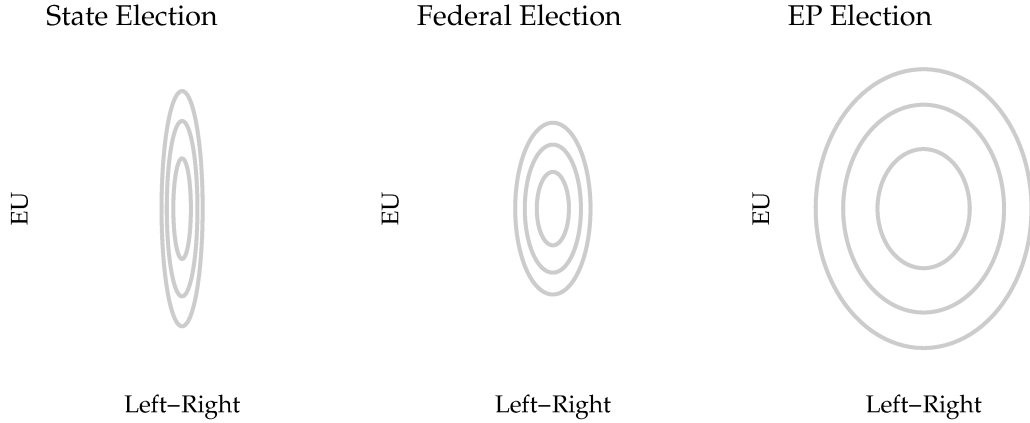
Figure 5.5: Utility indifference contours: *Observed Choice* model point estimates

Although the change in relative weights between elections is noticeable, the political significance of the results needs to be evaluated with caution. Even if we were to take the point estimates of the relative weights literally, the political consequences of the recalibration of the policy space between elections are likely to be minor in the grand scheme of things. Firstly, the findings indicate that left-right preferences play a substantial role also in the European arena. This continuity of left-right preferences can be counted as a major stabilizing factor for voting behavior across different elections. The change of relative weights likely needs to be more dramatic in order to motivate a large-scale realignment of voters in EP elections, i.e., EU issues need to take a dominant role over left-right preferences. Even if one would optimistically assume that measured distances on the Left-Right and EU dimension are on the same latent scale, the findings suggest that voters balance Left-Right and EU integration considerations only roughly 50:50. A second stabilizing factor is the configuration of party positions in the multidimensional policy space. Party positions on Left-Right and EU integration are in most party systems not independent. Instead, party positions tend to follow an inverted U-shape, with pro-EU taking centrist Left-Right positions and anti-EU parties located at the far left or right. As a consequence, most voters need to give up closeness on left-right in order to find a closer match on European issues. If voters weigh EU distances equally to Left-Right distances, the additional utility of voting for an alternative party due to more closeness on the EU dimension is at least partially, or even fully compensated by increased distance on left-right. While recalibration seems to be less consequential overall, there may however be specific parties that are particularly affected by recalibrating voters. These are parties that are challenged by alternative parties that represent differing view points on the EU dimension, but hold similar viewpoint on the Left-Right dimension.[18]

---

[18]The consequences of EU policy preferences for vote switching across elections is further

5.6 Conclusion

Second-order theory and the Europe Matters argument disagree about the role that European issues play in EP elections. While second-order theory sees voters as simply transferring their national-level decision making calculus to EP elections, the Europe matters argument builds on the idea that voters recalibrate their issue space in EP elections by putting a larger emphasis on European issues compared to national elections. In this Chapter, I tested these competing conceptions. I showed that the transfer and the recalibration hypotheses can be precisely formulated in the multidimensional model of spatial competition as diverging predictions about the change of relative dimensional weights across elections. I developed a research design that allows for the consistent estimation of the relative weights from panel data in the presence of unobserved heterogeneity. I find that the empirical evidence is much more consistent with the recalibration hypothesis than the transfer hypothesis. However, while my findings suggest that voters are more likely recalibrate than to transfer, the study fails to confidently reject the transfer hypothesis with the data at hand. Moreover, the findings suggest that the extent to which voters recalibrate is moderate at best: European issues do not play a dominant role in EP elections, as the traditional dimension of political conflict, Left-Right, maintains its relevance across all elections.

This study is novel in two major ways: First, it is the first study of EU multi-level voting behavior that seeks to go beyond merely establishing correlational patterns in observational data. By eliminating potential confounding due to unobserved heterogeneity, it presents a research design for exploiting panel data to the fullest which will enable researchers to place inferences on a much more solid footing. For the new panel data that will come online in the foreseeable future not only in the field of voting behavior research, this study may serve as a methodological blueprint. Secondly, it is, to my knowledge, also the first application of a conditional likelihood approach to multinomial choice data in political science. While the approach has its caveats, the proposed panel conditional multinomial logit model with time-varying parameters is show to be a powerful statistical tool that is particularly well suited to studying voter behavior across elections. It is particularly handy for voter panel surveys, which typically have only a small number of panel waves. Furthermore, the inclusion of alternative-specific variables is straightforward which makes it ideal for the estimation of empirical spatial models with panel data.

The study purposely focused on the internal validity of the estimates, and therefore necessarily neglects generalizability. The question of generalizability entails two aspects: First, how generalizable are the findings to the entire population on Bavarian voters in the given time-periods, and secondly how generalizable are they to other populations and time-periods. Both aspects depend on how one imagines

---

developed in Chapter 7.

true effect sizes to be distributed in the population or across countries and time. The first aspect follows from the fixed effects approach that only incorporates information from respondents who show varying voting behavior across elections (see Chapter 3 for a discussion). These respondents might be disproportionately drawn from a particularly mobile subset of the population who recalibrate more as they are also intellectually more flexible. In such a case, the obtained estimate is likely an overestimate of the average effect in the population. Therefore one should be careful to extrapolate the findings of this study to the entire population. However, I think a valid estimate for this group, even if not readily generalizable to the entire population, is more useful than an invalid estimate for the entire population. The second aspect, generalization to other countries and time periods, is probably more relevant for the comparativist literature on EP elections. Naturally, the findings of this study are not generalizable beyond the case of Bavaria 2013-2014. Spatial theory suggests that the extent of recalibration may indeed heavily dependent on party system characteristics and the particular electoral context. Speculation should be guided by plausibility considerations about how effect sizes are distributed across time and space. While it is clearly plausible that effect sizes vary substantially across settings, I think that it is less plausible the fundamental mechanism works differently in different settings: If Bavarians recalibrate, why shouldn't voters in other countries as well? Nevertheless, additional studies employing a similar research design in other electoral context are necessarily required to obtain a broader and deeper understanding of multi-level spatial voting in the European policy space.

5.7 Appendix to Chapter 5

Model 1: Observed Choice

Election	Variable	SPD	Greens	Free Voters	FDP	Left	Pirates	AfD
Federal	Constant	-0.247	-0.616	-1.676	-0.066	1.797	-0.587	0.624
		(0.194)	(0.243)	(0.264)	(0.236)	(0.455)	(0.48)	(0.274)
EP	Constant	0.402	0.628	-0.314	-0.514	2.698	-0.158	2.211
		(0.199)	(0.24)	(0.233)	(0.313)	(0.473)	(0.543)	(0.304)
State	EU distance			-0.001 (0.011)				
	LR distance			-0.029 (0.014)				
Federal	EU distance			-0.005 (0.01)				
	LR distance			-0.016 (0.013)				
EP	EU distance			-0.038 (0.011)				
	LR distance			-0.041 (0.013)				
Econom. situation		0.614	-0.56	0.06	-1.065	0.151	-0.335	-0.099
		(0.223)	(0.219)	(0.281)	(0.293)	(0.294)	(0.341)	(0.342)
Government satisf.		0.001	0.244	-1.045	0.68	-2.2	0.493	-0.28
		(0.336)	(0.446)	(0.355)	(0.765)	(0.712)	(0.302)	(0.268)

Table 5.2: Parameter estimates Model 1: Observed Choice

Model 2: Full Turnout (hyp.)

Election	Variable	SPD	Greens	Free Voters	FDP	Left	Pirates	AfD
Federal	Constant	-0.358	-0.717	-1.85	-0.143	1.309	-0.656	0.529
		(0.157)	(0.197)	(0.212)	(0.186)	(0.332)	(0.335)	(0.222)
EP	Constant	0.151	0.343	-0.714	-0.828	1.949	-0.299	1.8
		(0.159)	(0.192)	(0.187)	(0.245)	(0.342)	(0.361)	(0.233)
State	EU distance			0.004 (0.008)				
	LR distance			-0.029 (0.011)				
Federal	EU distance			-0.011 (0.008)				
	LR distance			-0.031 (0.010)				
EP	EU distance			-0.033 (0.008)				
	LR distance			-0.045 (0.011)				
Econom. situation		0.44	-0.71	-0.064	-1.114	0.231	-0.307	-0.28
		(0.177)	(0.188)	(0.229)	(0.248)	(0.243)	(0.262)	(0.257)
Government satisf.		-0.122	-0.012	-1.284	0.028	-1.743	-0.052	-0.742
		(0.274)	(0.347)	(0.323)	(0.411)	(0.47)	(0.248)	(0.231)

Table 5.3: Parameter estimates Model 2: Full Turnout (hyp.)

Chapter 6

Spatial voting with non-separable preferences

6.1 Multidimensional Spatial Voting with Non-separable Preferences (*with Lukas F. Stoetzer, published in Political Analysis 23(3)*)

Introduction

The spatial model of voting is *the* work-horse for theories and empirical models in many fields of political science research, such as the equilibrium analysis in mass elections (e.g., Schofield, 1978; McKelvey, 1986; Calvert, 1985; Lin et al., 1999), the estimation of legislators ideal points (e.g., Poole and Rosenthal, 1985; Clinton et al., 2004) and the study of voting behavior (e.g., Kedar, 2005; Dow and Endersby, 2004; Quinn et al., 1999; Alvarez and Nagler, 1998). Since Downs' (1957) seminal work, the theory has come a long way. Its generalization to the multidimensional policy space, the Weighted Euclidean Distance (WED) model (Davis et al., 1970; Enelow and Hinich, 1984b; Hinich and Munger, 1997) forms the stable theoretical foundation upon which nearly all present variations, extensions and applications of multidimensional spatial voting rest.

 While all these contributions have advanced spatial theory and methodology, we argue that an important concept that used to be an integral part of the multi-dimensional spatial model was somehow lost along the way: the idea that policy preferences on multiple dimensions may be non-separable. Non-separability means that utility derived from policy distance on one dimension is dependent on policy distances on other dimensions. Empirical as well as formal models commonly rely on *additive separable* specifications of the spatial utility function, which preclude this possibility as utility is only a function of the sum of dimension-specific policy distances. In separable specifications, the dimensions have "nothing to do with

each other" (Ordeshook, 1986, 90). In this paper we make the case that we should not stick to model specifications that preclude non-separability a priori. "There is nothing perverse about this preference rule" (Hinich and Munger, 1997, 86), and there are good reasons why real-world voter utility functions may be non-separable (Milyo, 2000).

The original mathematical formulation of the WED model explicitly allows for the possibility of non-separability, which is modeled as the product of dimension-specific directed distances (Davis et al., 1970; Enelow and Hinich, 1984b; Hinich and Munger, 1997). The textbook example for non-separability is a scenario of committee voting, where committee members vote sequentially on multiple issues (Hinich and Munger, 1997, 60). Here, preferences on one issue are conditional on the outcome of voting on another issue, if the two issues are non-separable. Non-separability has also been studied in similar contexts, such as legislative voting on multiple issues (Kadane, 1972; Kramer, 1972; Schwartz, 1977), voting in multiple simultaneous elections or referenda (Brams et al., 1997, 1998; Lacy and Niou, 2000), voting for multiple candidates (Cox, 1984; Benoit and Kornhauser, 1994; Lacy and Niou, 1998), in models of committee agenda control (Denzau and Mackay, 1981; Mackay and Weaver, 1981; Enelow and Hinich, 1984b) and EU council bargaining, where actors' spending preferences are conditional on expected policy outcomes (Finke, 2009; Finke and Fleig, 2013). A major contribution is also Lacy's model of survey responses, which explains item instability and question order effects by the non-separability of the underlying policy preferences (Lacy, 2001a,b).

The concept of non-separability has not yet been applied to the logic of multidimensional spatial voting in mass elections, in which voters choose policy platforms by evaluating their policy positions on multiple relevant policy dimensions. In the context of mass elections, non-separability means that a voter's evaluation of a platform on one policy dimension is conditional upon the position of this platform on other policy dimensions. If voters have non-separable utility functions, they no longer only evaluate platforms by their multidimensional distance from their ideal point, but also take into consideration how platforms combine directed distances over dimensions. As we show, these combinations, which we call policy packages, then have distinctive properties that voters care about. We suggest and provide evidence that accounting for non-separability might in fact be essential to our understanding of political choice. As all real-world policy platforms *only* come as policy packages, packaging might matter to voters.

Using the generalized Weighted Euclidean Distance model as the starting point, we discuss the theoretical foundations and implications of non-separability in mass elections. We show how non-separability can be incorporated and estimated in standard discrete choice models. In a Monte Carlo experiment we study the statistical consequences if the separability assumption is violated. We find that separable specifications then yield biased and/or unreliable estimates. In three empirical applications to national and presidential elections in the Netherlands,

the US and Germany we demonstrate that accounting for non-separability can lead to very different conclusions about the substantive role of policy preferences in explaining voting behavior. Lastly, we discuss how testing for non-separability should be an essential part of robustness testing in all empirical applications of multidimensional spatial models.

Non-separability in the Weighted Euclidean Distance Model

The canonical WED model (Davis et al., 1970; Enelow and Hinich, 1984b; Hinich and Munger, 1997) explicitly allows for non-separable voter utility functions. The spatial loss function for voter i and policy platforms j in a d-dimensional policy space is

$$U_{ij} = -\sqrt{[\boldsymbol{p_j} - \boldsymbol{v_i}]^T \boldsymbol{A}[\boldsymbol{p_j} - \boldsymbol{v_i}]} \tag{6.1}$$

where $\boldsymbol{v}$ is a coordinate vector of voter ideal point of length d, and $\boldsymbol{p}$ is a coordinate vector of policy platform positions of length d. $\boldsymbol{A}$ is a $d \times d$ weighting matrix.[1] Its diagonal entries are weights expressing the importance, or salience, voters attach to distances on the policy dimensions. Off-diagonal entries contain the separability terms. Preferences are separable iff $\boldsymbol{A}$ is a diagonal matrix, i.e., all off-diagonal entries are zero. $\boldsymbol{A}$ is subject to an important constraint: It is a symmetric positive definite matrix (Davis et al., 1970, 433).[2] Positive definiteness guarantees that the quadratic form $[\boldsymbol{p} - \boldsymbol{v}]^T \boldsymbol{A}[\boldsymbol{p} - \boldsymbol{v}]$ is positive for all $p_j - v_i \neq 0$. The substantive meaning of the symmetric property is that non-separability does not depend on which dimension is evaluated first. Positive non-separability parameters indicate a substitutional relationship between dimensional preferences, negative a complementary relationship.

Two directions of non-separability can be distinguished: dimensional preferences can be substitutes . Negative entries in the off-diagonals of $\boldsymbol{A}$ are associated with positive complementarity, and vice versa. This is due to the fact that distance enters utility negatively: If separability parameters are positive, larger distances on the individual dimensions lead to even higher loss in utility. As this nomenclature might be confusing at times, we will refer to distances on dimensions as substitutes if they are negative complementary, and complements if they are positive complementary.

Non-separability has far-reaching consequences for our understanding of spatial voting. Non-separability "requires that voters consider *all* issue positions before choosing *any*" (Hinich and Munger, 1997, 85). In effect, this means that voters

---

[1] $\boldsymbol{A}$ may be individual-specific or, as we assume here, homogeneous in the population. For a detailed discussion of the homogeneity assumption see Rivers (1988).

[2] A symmetric matrix is positive definite if all its eigenvalues are positive. A 2×2 matrix is positive definite if the product of the diagonal elements is larger than the product of the off-diagonal elements.
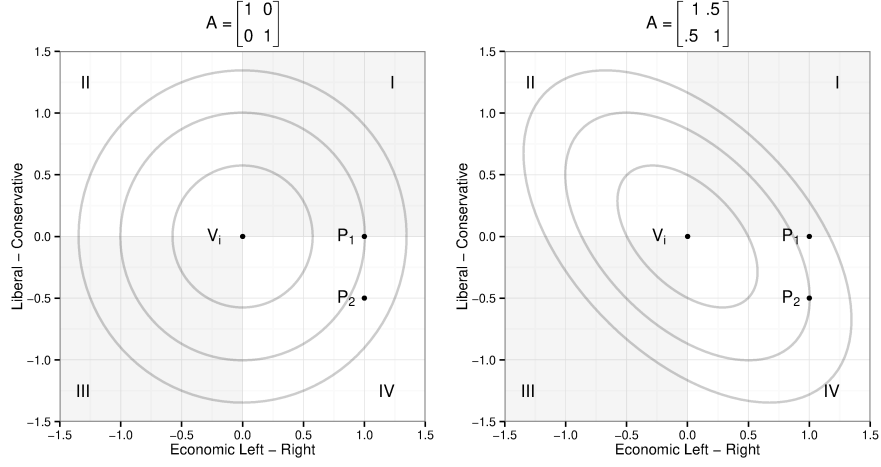
Figure 6.1: Exemplary choice scenario. Gray lines depict indifference contours. Left panel: Voter i with ideal point V and separable preferences prefers platform $P1$ over platform $P2$. Right panel: Voter i with non-separable preferences prefers $P2$ over $P1$.

evaluate policy packages and not the separate positions platforms take on each of the relevant policy dimensions. To illustrate the electoral consequences of non-separability, we confine our analysis to a two-dimensional policy space. Imagine the policy space to be defined by an economic left-right and a socio-cultural liberal-conservative dimension. Dropping the matrix notion,

$$U_{ij} = -\sqrt{a_{11}[p_{j1} - v_{i1}]^2 + a_{22}[p_{j2} - v_{i2}]^2 + 2a_{12}[p_{j1} - v_{i1}][p_{j2} - v_{i2}]}, \qquad (6.2)$$

where a_{11} and a_{22} are the dimension-specific salience parameters and $2a_{12}$ are the symmetric separability parameters.[3] Figure 6.1 depicts an exemplary spatial configuration in a Cartesian coordinate system, in which combinations of more leftist and more socially conservative positions are found in quadrant II, more rightist and conservative combinations in I, and so on.

When voter policy preferences are separable (left panel of Figure 6.1), i will prefer platform 1 over platform 2, as P_1 is located on a higher utility curve than P_2. i's preference ordering over platforms is reversed in the non-separable case, depicted in the right-hand panel of Figure 6.1. Here, platform 2 is at a higher utility curve even though both platforms are equidistant on the economic dimension, and platform 1 is congruent with i's liberal-conservative ideal point. The reversal of voter i's preference order over platforms is not a result of differences in proximity or salience, but due to the fact that the policy package offered by platform 2 yields higher utility than the package offered by platform 1. This property is inherent

---

[3]Note that assuming squared Euclidean metric, and $a_{12} = 0$, yields the commonly used specification of the spatial model: $U_{ij} = -a_{11}[p_{j1} - v_{i1}]^2 - a_{22}[p_{j2} - v_{i2}]^2$.

to the policy package. Platform 2 simply yields higher utility because it combines dimensional distances in a way that conforms with the direction of the voter's non-separability terms.[4]

If spatial preferences are non-separable, policy packages hold properties of their own. All real-world policy platforms only come as policy packages, and packaging, the way in which policy platforms combine policies, might matter to voters. We think there is no good reason why it should not. Thus taking non-separability into account has the potential of offering a more realistic picture of spatial voting. Whether policy preferences are non-separable is, at this stage, an empirical question.

A Conditional Logit Model with Non-Separable Preferences

McFadden's (1974) conditional logit is widely considered an appropriate discrete choice model to study spatial voting in multi-party systems (Alvarez and Nagler, 1998; Dow and Endersby, 2004).[5] Taking the conditional logit as our starting point, we propose a non-separable specification of the systematic component that follows from the WED model, and account for the positive definite constraint over $\boldsymbol{A}$. Unlike conventional specifications, which specify $\boldsymbol{A}$ as diagonal, we specify $\boldsymbol{A}$ to be symmetric and positive definite.

In the conditional logit, choice probabilities for voters $i \in (1, \ldots, n)$ choosing between policy platforms $j \in (1, \ldots, k)$ take the form

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{j=1}^{k} e^{V_{ij}}}. \tag{6.3}$$

V_{ij} is the systematic component of the voter utility function, which we specify as

$$V_{ij} = \theta_j + \boldsymbol{X_i}\boldsymbol{\delta_j} - \sqrt{[\boldsymbol{p}_j - \boldsymbol{v}_i]^T \boldsymbol{A}[\boldsymbol{p}_j - \boldsymbol{v}_i]}.^6 \tag{6.4}$$

---

[4]The consequences of non-separability may also be analyzed as sequential voting over individual dimensions. Keeping a party's positions on the economic dimension fixed, the voter ideal point on the liberal-conservative dimension shifts if preferences are non-separable. The new *conditional* ideal point is $v_2^*(p_1)$ is $v_2 - \frac{a_{12}}{a_{22}}(p_1 - v_1)$ (Enelow and Hinich, 1984b). Although conditional ideal points shift, this does not mean that *unconditional* voter ideal points are no longer fixed. Only the context changes. Voters still have an ideal package - their unconditional ideal point, but "there is no 'best' unique issue-by-issue ideal point" (Hinich and Munger, 1997, 61).

[5]Conditional logit, like multinomial logit, assumes the random error to be independently and identically distributed Type-1 extreme value. An undesirable feature of conditional logit is its reliance on the independence of irrelevant alternatives (IIA) (for a detailed discussion see e.g. Rivers, 1988; Alvarez and Nagler, 1998; Dow and Endersby, 2004). Multinomial probit has been considered as a solution. Specifying the systematic component in multinomial probit models as non-separable works the same way. Nevertheless, we opt for conditional logit because of its continuing popularity and since its computational convenience facilitates our Monte Carlo experiments.

[6]δ_j and θ_j are choice-specific parameters, while $\boldsymbol{A}$ is assumed to be homogeneous over choices

θ_j is a platform-specific constant that captures non-policy aspects, oftentimes labelled party or candidate valence. $\boldsymbol{X_i}\delta_j$ captures the effect of non-spatial individual-specific covariates on choice probabilities. The negative square root is the multidimensional spatial voting part as conceptualized in the WED model. $\boldsymbol{A}$ is a symmetric positive definite matrix. To incorporate this constraint in the maximum likelihood framework, we re-parameterize $\boldsymbol{A}$ as its Cholesky decomposition. This is a common procedure to solve numerically difficult optimization problems, such as the estimation of variance-covariance matrices (Pinheiro and Bates, 1996). $\boldsymbol{A}$ is parameterized as a lower triangular matrix $\boldsymbol{L}$, with $\boldsymbol{A} = \boldsymbol{L}^T\boldsymbol{L}$.

For a 2×2 $\boldsymbol{A}$ matrix, $\boldsymbol{L}$ contains three parameters.

$$\boldsymbol{L} = \begin{bmatrix} l_1 & 0 \\ l_{12} & l_2 \end{bmatrix} \tag{6.5}$$

$\boldsymbol{A}$ is restored post-estimation after maximizing likelihood with respect to $\boldsymbol{L}, \theta_k, \delta_k$.[7]

The consequences of misspecification: A Monte Carlo Experiment

Failing to account for non-separability if it is part of the true data generating process constitutes a misspecification of the functional form of how voter and party platform positions enter into the utility function. We study the consequences of misspecification using Monte Carlo methods. In order to obtain conservative estimates and to facilitate interpretation, we opt for a very basic design: Political choice in a policy space with two equally salient orthogonal policy dimensions.[8]

In separable specifications (6.6a), dimension-specific policy distances enter utility additively, in the non-separable specification (6.6b) they enter utility additively *and* multiplicatively .

$$U(v, p) = -\sqrt{a_{11}[p_1 - v_1]^2 + a_{22}[p_2 - v_2]^2} \tag{6.6a}$$

$$U^*(v, p) = -\sqrt{a_{11}[p_1 - v_1]^2 + a_{22}[p_2 - v_2]^2 + 2a_{12}[p_1 - v_1][p_2 - v_2]}. \tag{6.6b}$$

If the data generating process follows (6.6b), a conventional model (6.6a) is misspecified. Misspecification in choice models can result in biases that are anal-

---

and individuals.

[7]The likelihood function is given by the product over all realized probabilities. In order to identify this model, θ_k and δ_k are set to zero, for a baseline platform $j = k$. We use Broyden-Fletcher-Goldfarb-Shanno (BFGS) iterative numerical algorithm to maximize log-likelihood directly, using R's optim() function. In order to assure convergence on global maxima, maximization is repeated multiple times using randomly drawn starting values.

[8]There is no indication that the implications of our findings do not apply to higher-dimensional spaces as well. We suspect that the consequences of misspecification may become more pronounced as the number of dimensions increases and with larger differences in dimensional salience weights.

ogous to omitting important variables (Signorino and Yilmaz, 2003). Omitting non-separability terms when non-separability is present has the potential of leading to biased estimates of the salience parameters a_{11} and a_{22}, which express the importance of policy dimensions in the voters' choice calculus - the parameters of interest.

Under which circumstances are conventional estimates biased, in which direction and how severely? We show that the magnitude and direction of bias depends on the magnitude and direction of the non-separability parameter, and the distribution of platform positions in the policy space, relative to the distribution of voter ideal points. In order to test our intuition, we analyze the conditions under which the two expressions are not equivalent in expectation, i.e., $E[U(v,p)] \neq E[U^*(v,p)]$. If non-separability plays a role in the true data generating process ($a_{12} \neq 0$), the expressions are equivalent if $E([p_1 - v_1][p_2 - v_2]) = 0$. Without loss of generality, assume $E(v_{i1}) = E(v_{i2}) = 0$, which would be for example the case if voter ideal points are distributed independently multi-variate normal around the origin of a Cartesian coordinate system.[9] At this point, let us recall the properties of the variables v and p. While voter ideal points vary between voters, platform positions are fixed in a given sample. With $E(v)$ at $[0, 0]$, $E([p_1 - v_1][p_2 - v_2])$ can become either negative or positive, depending on p's position relative to $[0, 0]$. If $a_{12} > 0$, the multiplicative non-separability term is positive if the platform is in quadrant I or III, negative in quadrant II and IV, and vice versa if $a_{12} < 0$. The omitted term can therefore enter utility positively or negatively. As voters choose between multiple platforms, one has to consider the direction the omitted term has in expectation, over all platforms in the choice set. This is determined by the directed distance of platform positions relative to the expected voter ideal point. If platforms are positioned in a systematic way in the policy space relative to the expected voter ideal point, utility derived from policy distances under non-separability and separability rule is systematically different, if $a_{12} \neq 0$. The pattern of platform positions can be summarized by their correlation coefficient on the two dimensions. A positive correlation would indicate that positions along the first angle bisector of the Cartesian coordinate system are more likely, and negative along the second angle bisector. If platform positions are uncorrelated, positive and negative omitted terms cancel each other out in expectation. While conventional estimates would still be unbiased in expectation, one should expect that non-separability in this case increases the variance of the sampling distribution.

This rather intuitive analysis of the implications of violating the separability assumption motivates our design of a Monte Carlo experiment.[10] We study the

---

[9]Voter ideal points are also assumed to be uncorrelated over dimensions and platform positions on one dimension are independent of voter ideal points on the second dimension.

[10]An analytical solution is not easily tractable for the outlined choice model. Studies that are concerned with specification in choice models find omitted variable bias to be a more challenging problem than in the linear case (Yatchew and Griliches, 1985; Wooldridge, 2002). In probit models the estimates of a coefficient are generally biased downwards even if omitted variables

consequences of violating the separability assumption for the unbiasedness and sampling variance of the parameters of interest, the dimension-specific saliences, in three scenarios in which we vary the distribution of platforms in the policy space. In the first scenario, platform positions on the two dimensions are correlated. In the second scenario, platform positions on the two dimensions are uncorrelated, meaning that platforms are scattered unsystematically in the policy space. Platform positions are again heavily correlated in our third scenario, this time however negatively. For each of these three scenarios we draw 1000 voter ideal points and four party positions. Platform positions are drawn from a bivariate normal with variance terms of 0.5 and scenario-specific covariance terms.[11] Voter ideal points are drawn from a bivariate standard normal distribution. Voters choose between platforms according to the non-separable specification (Eq. 6.6b), assuming constant equal weights to both dimensions ($a_{11} = a_{22} = 1$) and varying degrees of separability. For each scenario, we vary the separability parameter a_{12} in 11 steps over the interval that meets the positive-definite constraint, $[-1, 1]$. 500 random samples for each combination of platform scenario and value of the non-separability parameter are drawn. For each of the resulting 16,500 unique datasets, we estimate a conventionally-specified separable (Eq. 6.6a) as well as a fully-specified non-separable model (Eq. 6.6b).

For each specific subscenario, we approximate the sampling distribution of the salience parameters by their empirical distribution in the 500 Monte Carlo samples. As we expect misspecification to render salience estimates inconsistent and/or inefficient, we report both the mean and 90% range of the bias in salience parameter estimates (Figure 6.2). The upper horizontal panel displays the additive separable model estimates, the lower horizontal panel the estimates obtained from the non-separable model. Vertical panels indicate the three main scenarios, in which platform positions were either correlated, not correlated, or negatively correlated. The upper horizontal panel shows that the misspecified model yields either biased and/or more unreliable estimates depending on the distribution of policy platforms in the policy space. If platform positions are positively correlated, and dimensional distances are complements ($a_{12} < 0$), the misspecified models underestimates the salience of both dimensions. If dimensions are substitutes ($a_{12} > 0$), their salience is considerably overestimated. The bias is not negligible: Even with moderately negative non-separability ($a_{12} = -0.6$), the salience parameters (a_{11} and a_{22}) are estimated at only around 50% of their true value, and are inflated by around 80%

---

are not correlated with other variables (Cramer, 2005). However, results from binary probit models do not straightforwardly carry over to unordered choice models. Lee (1980) explicitly studies omitted variable biases in the multinomial-logit context. His results of direction and strength of the bias are restricted to the case where omitted variables can be expressed as a linear function of other covariates with normal error. This is not the case here, since the omitted non-separability term can be expressed as a function of the distance terms. Moreover, the omitted variable bias is more complicated for the conditional logit model compared to the multinomial logit model.

[11]Correlations are set to .8, 0, and negative .8 respectively.
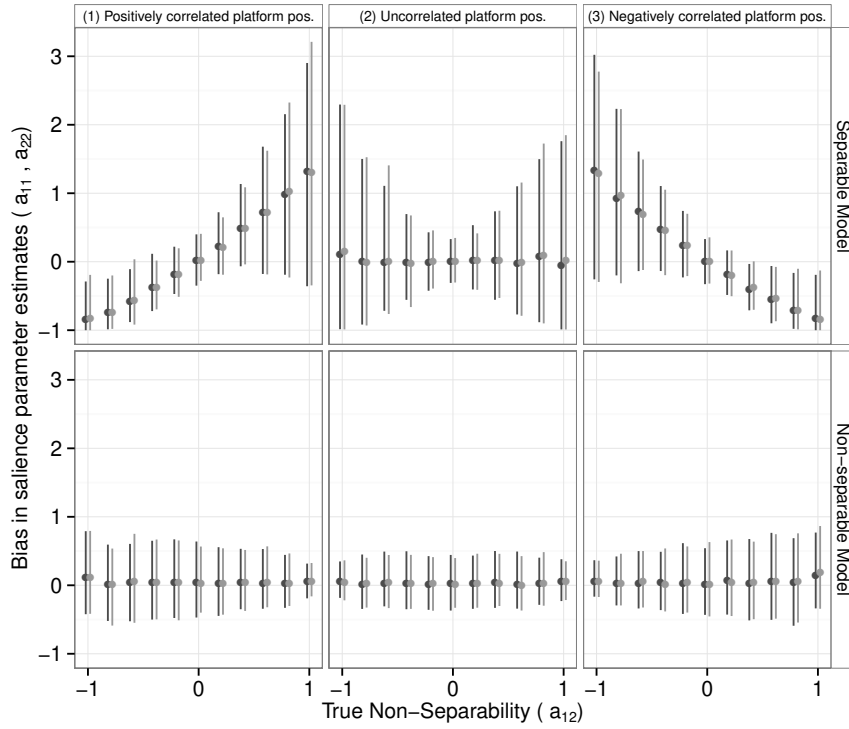
Figure 6.2: Monte Carlo Experiment: Bias in salience parameter estimates due to omitted non-separability. True salience parameters are both set at 1. Dots indicate mean estimates of salience parameters a_{11} (*dark gray*) and a_{22} (*light gray*). Vertical bars depict the 90% range of all estimates from Monte Carlo samples. (1) Platform positions randomly drawn from bivariate normal with positive covariance terms ($\rho = .8$), (2) Platform positions randomly drawn from bivariate normal with zero covariance terms ($\rho = 0$), (3) Platform positions randomly drawn from bivariate normal with negative covariance terms ($\rho = -.8$). Voter ideal points are drawn from a bivariate standard normal distribution. The non-separability parameter a_{12} is consistently estimated by the fully-specified model in all scenarios.

in the presence of moderately positive non-separability parameters (0.6). In case platform positions are negatively correlated, the direction of the bias is reversed, and positive non-separability parameters lead to a downward bias and negative to an upward bias. There is no theoretical bias if platform positions are uncorrelated, for all values of non-separability. However, the sampling variance increases considerably, as non-separability increases. This renders conventional estimates unreliable. As the lower horizontal panel indicates, a fully-specified model reliably recovers the true salience parameters in all scenarios.[12]. The non-separability parameter a_{12} is consistently estimated in all scenarios by the fully-specified model

---

[12]For a comparison of Root Mean Square Error (RMSE) and correctly predicted cases see the Supplementary Materials, which are available at http://dx.doi.org/10.7910/DVN/VCSRMX.

(see Supplementary Materials)

The message of our Monte Carlo experiment is clear: In the presence of non-separability, the statistical properties of a non-separable model are preferable to conventional, separable specifications. Dependent on party positions, a separable salience estimator is inconsistent and/or inefficient. The size and direction of the bias is dependent on the relative distribution of voters and policy platforms in the policy space. Even in a most basic case of two equally salient dimensions we find that bias can become severe and is not easily tractable. Given real-world data, it is therefore hardly ever apparent whether conventional models will run into problems and if they do, how severe these are.

In light of these insights, it is advisable to test the robustness of conventional models to potentially omitted non-separability. To demonstrate that non-separability is relevant not only in highly stylized Monte Carlo experiments but in explaining empirical phenomena, we proceed to three empirical applications in which we compare estimates obtained by separable and non-separable specifications.

Empirical Applications

Empirical studies of spatial voting in mass elections ultimately rely on estimates of voter ideal points and party platform positions. Voter ideal points are commonly inferred from voter surveys, which ask respondents to locate themselves on various policy or issue scales. Platform positions are inferred from either where respondents place policy platforms on these policy scales, or from outside sources such as expert surveys, roll call votes or analyses of platform manifestos. Over the years, a multitude of approaches has developed, each addressing some of the difficulties of estimating reliable ideal point and platform position estimates (e.g. Aldrich and McKelvey, 1977; Poole and Rosenthal, 1985; Kedar, 2005; Bartels, 2006; Jessee, 2009; Lo et al., 2013).

For our purposes we deem factor-analytic techniques which have been employed by Quinn et al. (1999), Schofield et al. (1998) and Schofield and Zakharov (2009) as most appropriate. Using expert and voter survey responses on multiple concrete issue scales, this approach allows for the placement of voters and platforms in a common multidimensional Euclidean policy space.[13] In effect, factor analytic methods approximate the structure of the policy space by analyzing the structure of voter survey responses. Policy dimensions are not defined a priori, but are rather uncovered using empirical data. The procedure can be described as follows: First, survey responses on a number of issue items are used to estimate an explanatory factor model. In these issue items respondents are asked to what degree they agree or disagree with specific policy statements, such as "Politics should abstain from intervening in the economy.". From the factor analytic solution, the num-

---

[13]We closely follow the procedure outlined in Quinn et al. (1999), we kindly ask the reader to refer to this article for details.

ber of underlying dimensions is determined that sufficiently describes the policy space. In all cases, we find a two-factor model to be most appropriate. We identify the first factor as an "economic left-right" dimension, and the second as capturing socio-cultural preferences, which we label as the "liberal-conservative" dimension.[14] Secondly, we conduct a two-factor confirmatory factor analysis (CFA) with uncorrelated factors. Using the CFA factor loadings we locate respondents in the policy space. In the third step policy platforms are projected into the same space, using the factor loading from the voter CFA. In two of our applications, positions of policy platforms on these issue items were not available. Here scales in expert surveys are identified that thematically match the policy scales of survey items as closely as possible. This necessarily involves a substantial degree of discretion, as survey items and expert scales are not identical (see Supplementary Materials).

We apply this method to three different empirical applications. The first uses the Euro-Barometer 11 dataset (Commission of the European Communities, 2012) on vote intention in 1979 in the Netherlands employed by Quinn et al. (1999).[15] The second application is concerned with voting in the 2008 US presidential election. Voter data is from the 10th wave of the American National Election Panel Study 2008-2009 (American National Election Studies, 2009), in which respondents were asked a battery of eight policy issue questions ranging from immigration to health care to taxation. Respondents were asked to locate themselves and the Democratic and Republican candidate on each of these 7-point scales, ranging from "strongly opposed" to "strongly in favor." After projecting both voter and candidate placements into the policy space, we average over the candidate positions to obtain a robust measure of candidate positions. The third application analyzes vote intention in the 2009 German federal election. Voter data is from the European Election Survey 2009 (EES) (van Egmond et al., 2013), which includes seven issue scales capturing attitudes towards immigration, extent of public services, state intervention in the economy, redistribution of wealth, criminal punishment and homosexuality. In order to locate party platforms on these scales, we identify seven issue scales in the Chapel-Hill Expert Survey 2011 (CHES) that match the EES scales. For a more detailed description of the question wording, highest density plot and factor loadings, see the Supplementary Materials.

---

[14]While this finding is hardly controversial for the Netherlands and Germany, where the two-dimensionality of the policy space is well established (see e.g. Benoit and Laver, 2006; Schofield et al., 1998), it may raise some eyebrows in the US application. Here the working consensus seems to be that inter-party political conflict is virtually unidimensional (McCarty et al., 2006; Poole and Rosenthal, 2007; Aldrich et al., 2014). However, unidimensionality of political conflict among political elites does not imply that voter policy preferences are unidimensional as well. Work on the structure of political ideology among Americans finds, very much in line with our findings, that voter preferences are structured by an economic and a social policy dimension (Shafer and Claggett, 1995; Treier and Hillygus, 2009; Klar, 2014).

[15]Replication data is well-documented and available online at `http://adm.wustl.edu/replication.php`. The model estimated in the original article however does not allow, unlike our WED model, for dimension-specific weights. Our results are therefore not readily comparable.

In accordance with standard model specifications (see e.g. Dow and Endersby, 2004; Kedar, 2005; Quinn et al., 1999), we include individual-specific control variables such as gender, age, education, religion, income or party identification. For each application we specify two vote choice models: a normal WED model that allows for dimension-specific weights, but assumes separability, and a non-separable WED model that allows for non-separability. We estimate the two models according to the conditional logit specification outlined above.

Results

Table 6.1 compares the spatial parameter estimates obtained from the separable and non-separable model specifications. We report estimated saliences of distances on the economic left-right dimension and on the liberal-conservative dimension, and the estimated separability parameter for the non-separable models.[16] As salience parameters are constrained to be positive, 95% confidence intervals are used to quantify estimation uncertainty.[17] How severe non-separability is in an estimated $\boldsymbol{A}$ matrix is not immediately obvious. We therefore report an intuitive measure of the degree of non-separability, that sufficiently summarizes both the direction and the degree of non-separability in two-dimensional policy spaces. This measure utilizes the positive-definite constraint to scale the separability parameter to the interval $[-1, 1]$, where -1 indicates perfect complements, and 1 perfect substitutes.[18]

Does modeling non-separability make a difference in the three applications we present here? If it does not, both model specifications should yield similar salience estimates. Salience estimates should not only have about the same magnitude, but should also not vary in their relative magnitude. Model fit is a second criterion. If non-separability is not an issue, the non-separable model should not exhibit a better model fit. As the models are nested, Likelihood Ratio tests are appropriate, which we report in the last row of Table 6.1. We also report expected Percentage Correctly Predicted (ePCP) (Herron, 1999) as an additional measure of predictive fit. In the first application, concerned with vote intention in 1979 in the Netherlands, the separable model suggests that only distances on the economical left-right dimension are relevant for vote choice. The coefficient of the second, liberal-conservative dimension is virtually zero, from which we would

---

[16]The paramter estimates of the individual-specific control variables are reported in the Supplementary Materials.

[17]Confidence intervals may be non-symmetrical due to the constraint induced by the Cholesky decomposition. The separability parameter is not subject to the positive constraint, is however constrained by the positive-definiteness of the $\boldsymbol{A}$ matrix.

[18]As $\boldsymbol{A}$ is a symmetric positive definite 2×2 matrix, the condition $a_{11} \cdot a_{22} - a_{12}^2 \geq 0$ holds. By rearranging we see that a_{12} is bounded between $\pm\sqrt{a_{11} \cdot a_{22}}$. Therefore $\frac{a_{12}}{\sqrt{a_{11} \cdot a_{22}}}$ is bounded between $[-1, 1]$. In order to convey the estimation uncertainty associated with the measure, we approximate 95% confidence intervals by calculating the degree of separability for repeated draws from the sampling distribution of $\boldsymbol{L}$.

Table 6.1: Empirical applications: Parameter estimates

DV: Vote choice	Netherlands 1979		United States 2008		Germany 2009	
	Sep.	Non-sep.	Sep.	Non-sep.	Sep.	Non-sep.
Economic Left-Right (a_{11})	0.73	0.92	0.81	0.44	0.8	0.74
	(0.43; 1.15)	(0.56; 1.38)	(0.42; 1.35)	(0.21; 1.07)	(0.38; 1.35)	(0.35; 1.46)
Liberal-Conservative (a_{22})	0	0.38	1.77	0.98	0.14	0.18
	(0; 0.2)	(0.19; 0.63)	(1.13; 2.56)	(0.43; 1.79)	(0.02; 0.41)	(0.04; 0.44)
Separability term (a_{12})	0	0.59	0	0.44	0	0.17
	.	(0.33; 0.93)	.	(0.19; 0.64)	.	(0.02; 0.38)
Degree of Separability	0	1	0	0.67	0	0.49
	.	(0.99; 1)	.	(0.19; 1)	.	(0.06; 0.85)
Number of Observations	529	529	1133	1133	619	619
ePCP	0.49	0.5	0.75	0.76	0.6	0.6
logLikelihood	−498	−483.5	−514.1	−510.9	−512.6	−510.6
Likelihood Ratio Test	6.7e-08		1.1e-02		0.04	

Note: Table reports point estimates and 95% confidence intervals in parentheses.

conclude that policy distance on this dimension is not associated with vote choice probabilities. This interpretation changes when we run a model that accounts for non-separability. Here we find distances on the second dimension to play a still subordinate, but noticeable role in explaining vote choices. The separability parameter estimate is significantly larger than zero. As the measure of the degree of separability indicates, policy distances on the two dimensions are estimated to be perfect substitutes. The Likelihood Ratio test indicates that the non-separable model fits the data considerably better. ePCP reveals a small increase in predictive accuracy when comparing the two models. In the application to voting in the US presidential election 2008, the separable model suggests that policy distance on the socio-cultural liberal-conservative dimension is more important to voters than distance on the economic dimension (95% C.I. for the difference in a_{11} and a_{22}: $[-1.69, -0.37]$). In contrast, the difference in coefficients becomes indistinguishable from zero in the non-separable model (95% C.I. $[-1.07, 0.09]$), indicating that policy distances on both dimensions play an equally important role in the vote choice mechanism. Again, the separability term is significantly larger than zero, as policy distances on the two dimensions are estimated to be partial substitutes. The degree to which preferences are substitutes is associated with substantial estimation uncertainty $[(0.19, 1]$. The Likelihood Ratio test indicates that the non-separable model fits the data better, ePCP shows a small increase in predictive accuracy. In the German case, non-separability seems to be less of an issue. Salience estimates are relatively robust to changes in the specification of the spatial model. The separability term has again a positive sign, but is only barely distinguishable from zero. The non-separable model fits the data only slightly better.
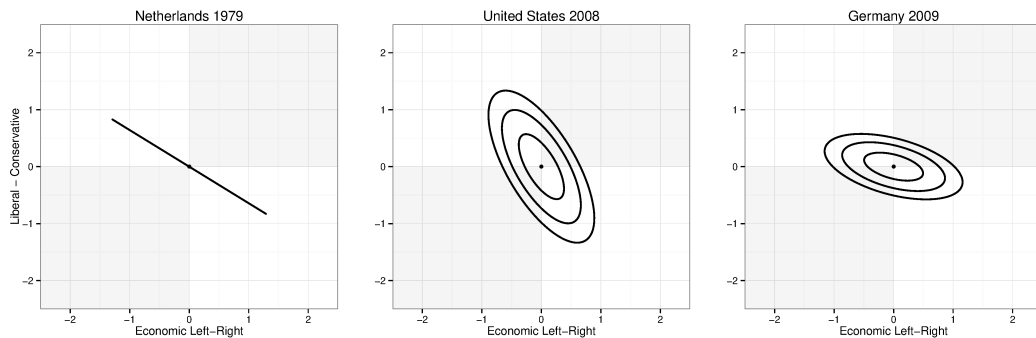
113

Figure 6.3: Indifference contours for A matrix point estimates.

Our findings indicate that non-separability makes a substantive difference for our understanding of voting behavior. In two out of three applications, allowing for non-separability has led to different conclusions about the role of dimensional distances in the voters' choice rationale. We find that in the Netherlands case, policy distances on the two dimensions are actually functional equivalents for voters. Figure 4 displays the shape of the indifference contours for point estimates of the spatial parameters. In the Netherlands case, where dimensional distances are perfect substitutes, the indifference ellipsoid practically collapses into a line: Even though voters hold meaningful preferences on two dimensions, preferences are translated into vote choices using a one-dimensional concept. Effectively, non-separability leads to a reduction in the dimensional complexity of political choice. In the application to voting in US presidential elections, separable models would have found that voters assign a higher weight to preferences on the liberal-conservative dimension than to the economic dimension. Preferences on the two dimensions have about the same importance to American voters, if we allow for non-separability. Although the degree of non-separability varies considerably, our results are remarkably consistent. Preferences on the two dimensions under investigation are substitutes in all three electoral contexts we have analyzed. A voter located at the center of the policy space, choosing among two parties which take more rightist position than the voter on the economic left-right dimension prefers the party that is more liberal on the cultural dimension to one that holds his preferred position on the cultural dimension. This is because the more liberal party compensates distance on the economic dimension by distance in the opposite direction on the cultural dimension.

Conclusion

In this paper we have advocated to bring back a concept that once used to be an integral part of the multidimensional spatial framework: Multidimensional spatial preferences may not be independent of each other, they may be non-separable. In mass elections, non-separability means that voters do not separately evaluate

the positions policy platforms take on multiple dimensions. They rather care about the policy packages that platforms offer. If the way in which platforms combine positions matters to voters, we need to allow for non-separability in our models. We present a consistent way of including non-separability in vote choice models. We find that failing to allow for non-separability can seriously undermine the validity of empirical tests of spatial theory. Our Monte Carlo experiments show that conventional salience estimates are biased and/or unreliable in the presence of non-separability. The magnitude and direction of bias depends on a non-trivial interaction between the degree and direction of non-separability and the distribution of policy platforms relative to voter ideal points in the policy space. If faced with real-world data, it is therefore not apparent whether omitted non-separability might be problematic in a statistical sense. Even more so when voter preferences are not limited to a low number of latent policy dimensions, but are defined over a high-dimensional issue space, where we might reasonably suspect non-separability to be the rule rather than the exception. Thus, to be on the safe side, careful researchers should test the robustness of obtained estimates with non-separable model specifications. Researchers who want to rely on linear predictor functions fit for out-of-the-can statistical programs can accommodate for non-separability by adding the products of all combinations of *directed* dimension-specific distances to the systematic component. In a two-dimensional policy space and using squared Euclidean metric, voter utility can be specified as $U_{ij} = \beta_1(p_{j1} - v_{i1})^2 + \beta_2(p_{j2} - v_{i2})^2 + \beta_3(p_{j1} - v_{i1}) \times (p_{j2} - v_{i2})$. β_3 then can be interpreted as the non-separability parameter. If $\hat{\beta}_3$ is significantly different from zero, non-separability is an issue.

On a more positive note, addressing the issue of non-separability can help reduce bias and/or increase the precision of spatial estimates. Our findings are therefore potentially relevant for all empirical applications of multidimensional spatial voting models that base their inferences on spatial salience estimates. Most prominently in studies of voting behavior that compare the importance voters assign to various issues or dimensions, such as the question whether economic issues trump "moral" issues or vice-versa in U.S. presidential elections (Bartels, 2006; Gelman, 2008), how party system compactness relates to relative issue importance (Alvarez and Nagler, 2004), or which role attitudes towards Europe play in explaining electoral behavior in European elections, relative to left-right preferences (de Vries et al., 2011; Hobolt et al., 2009; Lo et al., 2013). Here, the validity of spatial estimates, and the conclusion that one dimension is relevant, not relevant, or more relevant for political choice than other dimensions may depend on whether voter utility functions are specified as separable or non-separable.

Apart from being a safeguard against statistical pitfalls, caring about non-separability opens up new interesting perspectives on the structure of voter preferences in the multidimensional policy space. Analogous to a resource or budget constraint (Milyo, 2000), non-separability in mass elections can be imagined as an *ideological* constraint (Converse, 1964). The constraint determines which policy

packages are more attractive to voters. In all of our empirical applications we find a substitutional relationship between economic and social policy preferences. Such a relationship would indicate that the two policy dimensions share at least to some degree the same function and fulfill the same voter needs. If preferences are perfect substitutes, it becomes hard to argue that voters really care about individual policy dimensions. Although voters have well-defined preferences on these policy dimensions, what they really care about when choosing representatives is a lower-dimensional concept such as a single ideological dimension. In such a case, non-separability leads to a reduction in the effective dimensionality of the policy space, linking multidimensional mass preferences to unidimensional inter-party competition at the elite level.

Non-separability therefore not only informs the study of voting behavior but also the formal analysis of party competition. How do equilibrium configurations change if parties maneuver in a multidimensional policy space with non-separable voter preferences (Merrill III and Adams, 2001; Schofield and Sened, 2005), and can non-separability explain the empirical phenomenon that in many countries parties align along a single axis of competition even though the policy space is two-dimensional (Shikano, 2008)? Non-separability can also be brought to bear on the empirical question whether political polarization in the US has increased in recent decades (Fiorina et al., 2008; Levendusky, 2009; Aldrich et al., 2014). Polarization may express itself not only in changes in voter preferences, issue partisanship, or issue alignment, but also in increasing non-separability of voter utility functions.

The implications of non-separability are multifaceted. Additional research is required to deepen our understanding of non-separability, and to thereby deepen our understanding of spatial voting in multidimensional spaces. To be sure, we do not suggest that multidimensional representations of voter preferences are generally preferable to one-dimensional representations. But we argue that if empirical researchers opt for multidimensional spatial representations, the potential non-separability of spatial preferences needs to be addressed. We hope that the findings and methods presented in this paper can serve as a guideline for future research to bring non-separability back into the fold of spatial theory.

6.2 Testing for non-separability of Left-Right and EU preferences

In Stoetzer and Zittlau (2015), we argue that non-separability should be treated as an integral part of the spatial model whenever the policy space is multidimensional. We also show that omitted non-separability may have detrimental statistical consequences for the integrity of dimensional weight estimates. Clearly, these insights are pressingly relevant for the investigation of relative dimensional weights of the EU dimension that I presented in Chapter 5. Accordingly, this section replicates the analysis using the analytical toolkit that we developed in Stoetzer and Zitt-

lau (2015). The re-analysis serves two purposes: Firstly, it tests the robustness of the results to potential violations of the separability assumption. Secondly, it answers the substantively interesting question whether Left-Right and EU integration preferences are separable or non-separable. The study of non-separability of preferences in the European policy space is novel in its own right, and updates the EU issue voting model to the latest developments in the spatial voting literature.

As described in the Conclusion section of Stoetzer and Zittlau (2015), including non-separability in standard estimators of spatial voting models is relatively straightforward. If the distance metric is squared Euclidean, (symmetric) non-separability is captured by simply including the product of directed distances as an additional independent variable. Additionally, the positive-definite constraint on the A matrix can be dropped to simplify estimation further. As the proposed modification concerns the linear predictor function, this strategy is not only viable for the MNL model as used in our article, but applies to PML as well. My non-separability estimator therefore differs only slightly from the the setup I have developed in the previous Chapter. I only change the linear predictor function for the spatial component to $\beta_{eu_t}(p_{eu_{jt}} - v_{eu_{it}})^2 + \beta_{lr_t}(p_{lr_{jt}} - v_{lr_{it}})^2 + \beta_{nonsep_t}(p_{eu_{jt}} - v_{eu_{it}}) \times (p_{lr_{jt}} - v_{lr_{it}})$. This means that three additional parameters are estimated (β_{nonsep_t}), which express the direction and magnitude of the non-separability in the election-specific spatial voting calculus of respondents.

Election	Dimension	Observed Choice		Full Turnout (hyp.)	
		β	s.e.	β	s.e.
State	EU	-0.002	(0.011)	0.003	(0.008)
	LR	-0.03*	(0.014)	-0.03*	(0.011)
	nonsep	-0.022	(0.017)	-0.019	(0.014)
Federal	EU	-0.005	(0.01)	-0.012	(0.008)
	LR	-0.016	(0.013)	-0.032*	(0.01)
	nonsep	-0.011	(0.016)	-0.002	(0.013)
EP	EU	-0.04*	(0.011)	-0.035*	(0.009)
	LR	-0.045*	(0.014)	-0.049*	(0.011)
	nonsep	-0.016	(0.017)	-0.014	(0.014)

Note: Table reports point estimates and standard errors in parenthesis. $^*: p < .05$

Table 6.2: Non-separability model: Spatial parameter estimates

Taking a first cursory look on Table 6.2 that presents the spatial parameter estimates, one can see that the substantive findings remain largely the same. While only the weight of Left-Right distances is distinguishable from zero in the State and

Federal election (with the exception of LR in the Federal election on the Observed Choice data set), distances on the EU dimension do seem to play a role in the EP election as well. Investigating the additional non-separability parameter estimates, one can see that these are indistinguishable from zero in all elections. This means the null hypothesis that Left-Right and EU preferences are separable can not be rejected on the basis of my data.

If one would take the point estimates literally, the negative sign of the non-separability indicates that preferences on the two dimensions act as substitutes rather than as complement. If distances point in the same directions, e.g., a party is more rightist and more pro-EU that the voter, the utility of voting for said party is more negative compared to an equally distant party that is located in the direction more right/anti-EU or left/pro-EU. Putting it the other way around, voters are more forgiving to parties that align along the left/pro-EU and right/anti-EU diagonal, as seen from the voter's ideal point. This can also be seen in Figure 6.2 that depicts the estimated shape of the spatial utility indifference contours. In the State and Federal election, the utility indifference ellipsoid practically collapses into a line along which parties more or less yield the same utility. As a result, the policy space becomes one-dimensional, orthogonal to the utility indifference "line". The line is close to orthogonal to the Left-Right dimension in the State election, which means that voters are almost exclusively concerned with distances on the Left-Right dimension when they make up their mind for whom the vote for. In the Federal election, the line is more tilted, indicating the European integration aspects may play a small role, but are largely considered to be functionally equivalent to Left-Right aspects by voters. While the voter utility function is largely consistent with a uni-dimensional Left-Right policy space in the State and Federal election, the policy space clearly become more two-dimensional and mostly separable in the EP election, indicated by the almost circular utility indifference contours. European integration aspects seems to be evaluated in their own right, and become consequential for vote choice.

Although the analysis of the shape of utility function gives interesting insights into the configuration of the European policy space, they have to be taken with a grain of salt. These are merely point estimates - there is large uncertainty about the non-separability parameter. Indeed there is barely any evidence for non-separability of Left-Right and EU preferences. This is good news for students of European voting behavior, and in fact for this dissertation. Describing the relevant policy space for voters in EP elections as consisting of a separable Left-Right and EU dimension is most likely accurate. Also the analysis has revealed that the findings of the previous chapter are robust to a lifting of the separability assumption. A very similar picture arises: While spatial voting in State and Federal elections is dominantly structured by Left-Right preferences, EU preferences seem to become more important in EP elections. However, the findings do not suggest that EU preferences play a dominant role. Nevertheless, these patterns are consistent with the predictions of the recalibration hypothesis that sees voters as

State Election   Federal Election   EP Election

– EU +   – EU +   – EU +

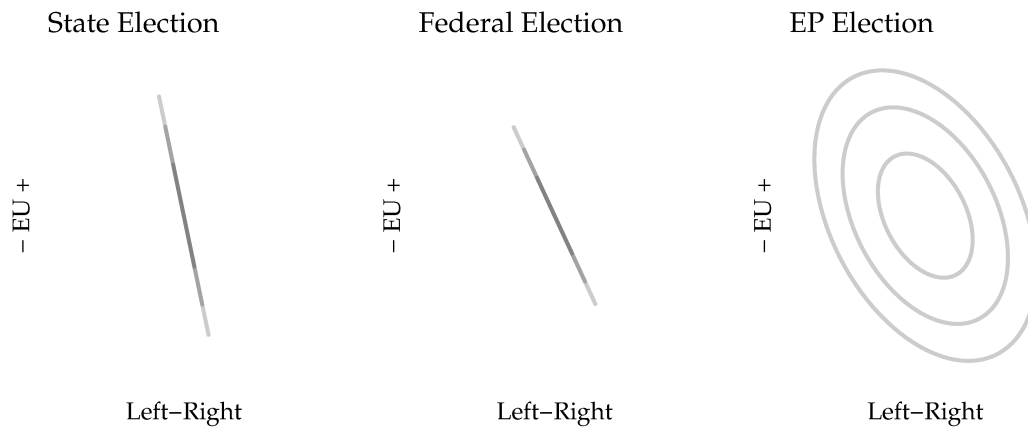Left–Right   Left–Right   Left–Right

Figure 6.4: Utility indifference contours: *Observed Choice* model point estimates

changing the relative weight of policy considerations according to their relevance
at different levels of government.

Chapter 7

EU integration preferences and vote switching in EP elections

7.1 Introduction

The question why some voters vote for different parties in European Parliamentary (EP) elections than in first-order elections (FOE) has been at the core of the literature on European multi-level democracy since the first EP election in 1979. The traditional and dominant theoretical lens through which scholars, media and political actors alike view and interpret EP elections, the second-order elections model (Reif and Schmitt, 1980) argues that the key factor that explains why voters vote for different parties is that there is simply "less at stake" in EP elections. Second-order theory specifically precludes the possibility that voters switch their vote to a different party because they perceive that there are different issues at stake in EP elections, on which other parties may better represent the voter. Instead, voters are thought to simply "apply their evaluations of national-level phenomena to the EU level", which has been labeled "the transfer hypothesis" (Clark and Rohrschneider, 2009, 645). Opposition to this interpretation has been voiced in the form of the Europe Matters argument. Several contributions have argued that due to the increasing powers of the European Parliament, EP elections have become more important and more European after all (e.g., Hobolt et al., 2009; Hobolt and Wittrock, 2011; de Vries et al., 2011). The key proposition of this alternative strand of research is that voters vote for different parties in EP elections because they perceive the election to be about different issues. Particularly, voters recalibrate their issue space by putting a larger emphasis on European issues when they decide whom to vote for in EP elections. In Chapter 5 and 6, this dissertation has presented empirical evidence that voters indeed recalibrate their issue priorities between FOE and EP elections. However, these chapters have remained mostly quiet about the political consequences of this recalibration. An important research question therefore remains to be answered in the context of this dissertation: To

what extent do preferences on European issues motivate vote switching in EP elections?

As has been discussed in Chapter 3, previous studies of vote switching in EP elections have faced major theoretical and methodological shortcomings. I have argued that previous studies developed their hypotheses based on an incomplete spatial model. Common practice is to collapse different vote choices across FOE and EP elections into a binary variable, and then to establish to which degree policy disagreement on European issues with the party chosen in the FOE explains this variable. I find that this practice misses a fundamental point about the process of voting in two successive elections. For each voter, there is a party of origin, the party chosen in the FOE, and a party of destination, the party chosen in the EP election. If the party of origin is not the same as the party of origin, the voting behavior can be qualified as vote switching. Modeling vote switching as a binary variable however means loosing the ability to take characteristics of both the party of origin and destination into account. This is necessarily required if one argues based on a spatial model of voting in multi-party systems. In the spatial model, the characteristics of the party of destination, and in fact the characteristics of all possible parties of destination, matter for vote switching.

This chapter proposes an alternative way to study vote switching. I argue that vote switching is best represented as a Markov process, where the transition probabilities, the probability of a change from one state to a second state are expressed as a function of policy preferences. This allows for modeling vote switching probabilities in line with spatial theory - conditional of the policy positions of all possible parties of destination. I propose a multinomial transition model to estimate the impact of policy distances on a Left-Right and EU integration dimension on transition probabilities. Based on the estimates, I investigate the consequences of policy disagreement on European issues for vote switching in EP elections using simulation techniques.

7.2 Descriptive Analysis

Given vote choice data, a frequency cross-table of vote choices in the Federal and EP election summarizes the observed voting sequences. Table 7.1 presents the (post-election) reported vote choices in my data. Vote choices in the previous Federal election are arranged horizontally, and EP election vote choices vertically. The frequencies on the diagonal of the table signify the loyal voters that stayed with the party they voted for in the first election, off-diagonal entries are inter-party movements signifying vote switching behavior. First of all, one can see that the amount of vote switching in the data is substantial. Roughly one third of the voters that participated in both elections voted for a different party in the EP election. At this point, binary switching models collapse the frequency table into a binary variable: All observations on the diagonal are coded as partisans,

all observations on the off-diagonals are coded as switchers. Obviously this means a drastic loss of information. Information is lost because all parties of origin and destination are effectively treated alike.

EP vote	Federal election party vote								
	AfD	CSU	FDP	Free V.	Greens	Left	Pirates	SPD	
AfD	92	41	10	4	5	12	1	24	189
CSU	10	485	35	8	5	7	6	20	576
FDP	2	15	22	3	0	0	0	3	45
Free Voters	7	28	2	29	4	1	2	19	92
Greens	1	35	5	5	85	5	5	35	176
Left	1	6	0	2	2	27	0	19	57
Pirates	4	1	1	0	1	0	11	4	22
SPD	6	61	8	4	25	13	2	204	323
	123	672	83	55	127	65	27	328	1,480

Table 7.1: MEDW Bavaria panel: Vote choice frequency table

Instead, I propose to investigate the transitions not as binary (switch vs no switch), but as a transition process from a party of origin to a party of destination. If the party of destination is not the same as the party of origin, a transition can be classified as a switch. Such a sequence of choices constitutes a first-order Markov process. First-order Markov transition probabilities, the probability of voting for party x in election 2 conditional on having voted for party y in election 1, can be estimated from this frequency table. The maximum likelihood estimate of the transition probability is the sample conditional probability that a vote for party x will follow a vote for party y (Anderson and Goodman, 1957). Therefore, transition probabilities are simply the column percentages of Table 7.1, the resulting Table 7.2 is the transition matrix collecting all transition probabilities. E.g., the probability of voting for the AfD in the EP election, conditional on having voted AfD in the previous Federal election is .75 etc.

EP vote	Federal election party vote							
	AfD	CSU	FDP	Free V.	Greens	Left	Pirates	SPD
AfD	0.75	0.06	0.12	0.07	0.04	0.18	0.04	0.07
CSU	0.08	0.72	0.42	0.15	0.04	0.11	0.22	0.06
FDP	0.02	0.02	0.27	0.05	0	0	0	0.01
Free Voters	0.06	0.04	0.02	0.53	0.03	0.02	0.07	0.06
Greens	0.01	0.05	0.06	0.09	0.67	0.08	0.19	0.11
Left	0.01	0.01	0	0.04	0.02	0.42	0	0.06
Pirates	0.03	0.001	0.01	0	0.01	0	0.41	0.01
SPD	0.05	0.09	0.10	0.07	0.20	0.20	0.07	0.62

Table 7.2: MEDW Bavaria panel: Sample Transition matrix

From the sample transition matrix one can already see some notable patterns in the data. First of all, parties differed substantially in their ability to retain former voters in the EP election. Retention rates are the entries on the diagonal of the transition matrix, the switching rate is obviously one minus the retention rate. AfD and CSU have the highest retention rates (.75 and .72), followed by Greens and SPD (.67 and .62). The very low retention rate of the FDP (.27) requires specific attention. First of all, it reflects the breakdown of the FDP's electoral support after loosing parliamentary representation after the 2013 Federal election. An additional plausible explanation is that a substantial part of the FDP votes in the Federal election were rental votes from CSU supporters (Gschwend et al., 2016), who returned to their party in the EP election.

However, one is usually not interested in the sample transition matrix itself. What I am interested in is the transition matrix conditional on voter characteristics, such as the EU integration preferences of voters. One simple way of doing this is to investigate the sample transition matrices for subsets of voters with different levels of EU integration support. In effect, I calculate the transition matrices for subsets of the sample, each subset containing only voters with similar EU policy preferences. As described in Chapter 3, I use the EU integration self-placement on a 11-pt scale. Figure 7.1 graphically displays the transition probabilities of voting for a different party, the switching rate, for selected parties, conditional on the respondents' self-placement on the EU integration scale. I omit smaller parties as the low number of respondents that voted for these parties make a sensible interpretation of transition probabilities in the subsets impossible. Furthermore, I also omit all frequency matrix cells with less than five observations. Mean perceived party positions on the EU integration dimension are signified by the dashed vertical lines. Point sizes and fit weights are proportional to the number of respondents in the respective cell of the subsetted frequency matrix.

The observed switching rates of the parties under investigation vary substantially, dependent on the EU integration preferences of the voters. Switching rates for the anti-EU party AfD (EU integration position: 1.5) seem to increase for former AfD voters that hold more pro-EU positions than the party. The most loyal AfD voters, with a switching rate of only around 10 percent, are those that hold more extreme anti-EU preferences than the mean perceived party position. Former CSU voters seem to be most loyal when they hold centrist positions, which conforms with the centrist CSU position on EU integration (5). Former CSU voters that hold more pro-EU integration positions than their party have an only slightly higher chance to switch. Observed switching rates increase more steeply for former CSU voters who hold more anti-EU preferences. For the pro-EU Green party (6.5), more pro-EU voter preferences seem to have a positive monotonic effect on the probability of former Green voters to cast a vote for the Green party in the EP election as well. The pro-EU party SPD seems to have a relatively stable switching rate, only very EU-skeptic former SPD voters are more likely to desert their party.
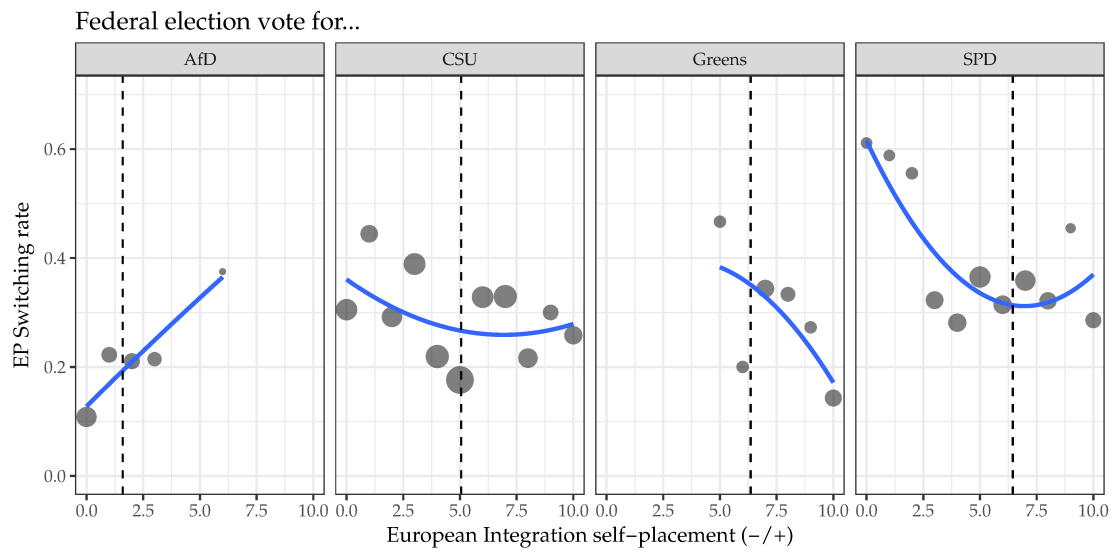
Figure 7.1: EP Switching rates, conditional on EU integration preferences at the point of the EP election. 2nd-order polynomial line of best fit.
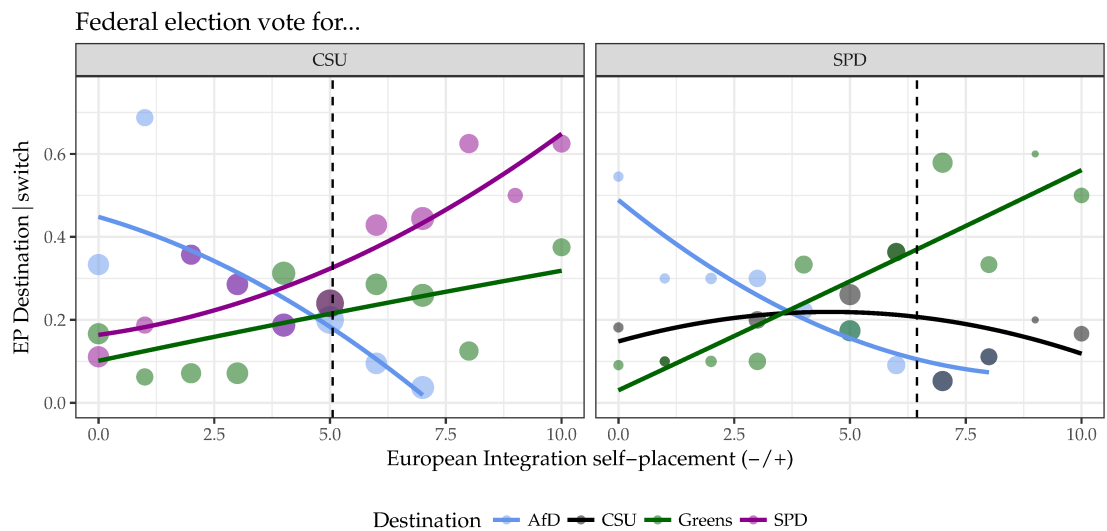

Figure 7.2: Probability of destination given a switch, conditional on EU integration preferences at the point of the EP election. 2nd-order polynomial line of best fit.
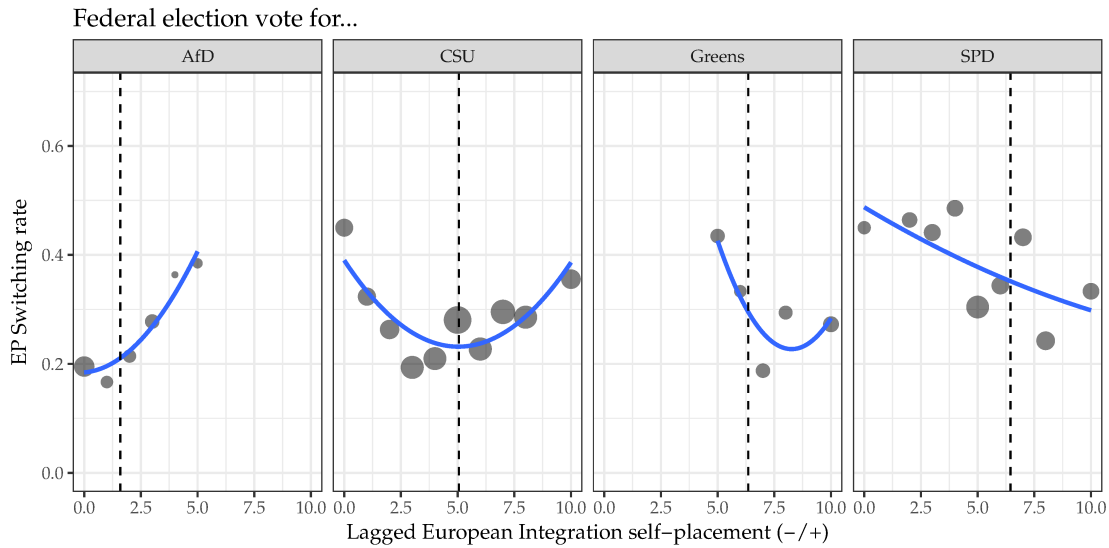
Figure 7.3: EP Switching rates, conditional on EU integration preferences at the point of the Federal election. 2nd-order polynomial line of best fit.

A valid point of criticism about this investigation would be that the vote switchers might be merely voters that have updated their EU integration preferences since the point in time when the Federal election took place. Then vote switching would be the consequence of a change in preferences in the mean time rather than due to the change in electoral context. Fortunately, this alternative explanation for the patterns can be tested with the panel data at hand. Instead of subsetting based on EU integration preferences reported at the time of the EP election, I subset on lagged preferences, i.e., the preferences that the respondents reported at the time of the Federal election. If the switching patterns are not (only) due to change in preferences, similar patterns should be observed when using these lagged preferences. As can be inferred from Figure 7.3, the previously observed association between voter EU integration preferences and switching rates remains substantively the same for all parties under investigation. This may be counted as clear evidence that something more fundamental is going on than simply voters changing their preferences between elections. Especially noteworthy is the pattern of CSU switching rates, which neatly follows an inverted U-shape. Apparently, lagged preferences seem to be a better predictor for EP voting behavior than EU integration preferences at the point of the EP election in predicting vote switching of CSU voters.

The bivariate investigation establishes clearly that policy distance on EU integration and vote switching in EP elections are related. However, already at this early stage of the analysis it is to be remarked that the observed patterns are not in line with the hypotheses developed in previous studies (Hobolt et al., 2009, see also Chapter 2). For at least two of the investigated parties it is not the case
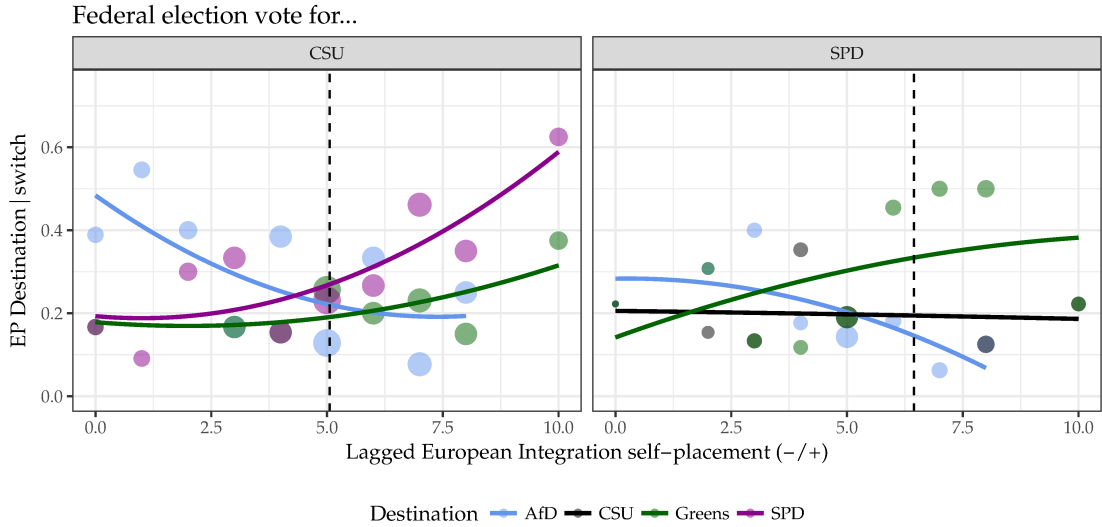
Figure 7.4: Probability of destination given a switch, conditional on EU integration preferences at the point of the Federal election. 2nd-order polynomial line of best fit.

that a voter's policy distance from her party of origin's position on EU integration increases her probability of vote switching. The most loyal AfD and Green voters seem to be those who hold more extreme positions on EU integration than their party. For the two largest parties, which were also both part of the Federal government coalition, the observed patterns suggest that only deviations in one direction, towards more anti-EU integration preferences, have a noticeable impact on the likelihood to switch. My critics might argue that measurement issues, such as different scale perceptions, drive the observed patters. However, I think it would not be a fair argument to try to dismiss the data, as there is a more sparse explanation for the discussed patterns. In fact, the patterns are fully consistent with a two-dimensional model of spatial voting. If both distances on EU integration and on Left-Right are relevant for voters in EP elections, as I have shown in the previous Chapter and will show in the following analysis, not only policy distance to the party of origin, but multidimensional policy distance to all available parties of destination determines vote choice probabilities. Based on spatial theory, former AfD voters that are more extreme on EU integration than their party are indeed expected to be more likely to vote for the AfD again, since the AfD is by far the most anti-EU party of all available parties. The same is true, although in the other direction, for extreme pro-EU Green voters. The spatial explanation for the CSU and SPD patterns is a bit more subtle. Former CSU voters with anti-EU preferences have a viable alternative in the AfD, whose Left-Right position is similar to the CSU. For more pro-EU CSU voters, there are pro-EU integration alternatives in the Greens and SPD, however these alternatives are also more leftist, meaning additional closeness on EU integration may be partially offset by additional dis-

tance on Left-Right. The effect of policy disagreement should therefore be more pronounced in the anti-EU direction. For anti-EU SPD voters, viable alternatives are located to the left (Left party), and right (AfD), which means that policy disagreement on EU integration needs to be substantial to offset the switching costs due to increased distance on Left-Right.

7.3 Modeling transition probabilities

The descriptive investigation of transition probabilities suggests that vote switching in EP elections may indeed be driven by policy considerations on EU-related issues. However, a bivariate investigation can be misleading since other factors might be confounding the results. I already identified one confounding factor in the above discussion. The proposed spatial voting model posits that vote choice is a function of the policy preferences not only on EU issues, but on left-right issues as well. While it is theoretically possible to extend the descriptive approach above by further subsetting on left-right preferences, in practice this is not a viable strategy due to data constraints. Very quickly there are simply not enough data points left in most cells of the transition frequency tables to make reliable inferences. Also other factors might act as confounders. E.g., voter in larger policy disagreement with the party they voted for in the Federal elections might be more likely to lack a strong party identification, which in turn makes them more likely to switch.

The analytical solution I shall pursue in this paper is to employ a parametric model that allows for the expression of transition probabilities as a function of multiple voter characteristics. Fortunately, it is fairly straightforward to modify conventional logit models to fit the structure of a Markov process. Cox (1970, 72ff) was among the first to note the link between Markov transition probabilities and the logistic regression framework. A logit model for transition probabilities is generally called a (Markov) transition model. Transition models belong to the standard statistical repertoire in many disciplines, such as bio-statistics (e.g., Diggle et al., 1994), econometrics (e.g., Boskin and Nold, 1975) and sociology (e.g., Coleman, 1964), but have found their way into the political science mainstream only fairly recently (Epstein et al., 2006; Jackman, 2000; Hillygus and Jackman, 2003; Hillygus, 2007). In its general form, transition models are usually written as $logit[Pr(y_{it} = 1|y_{i,t-1})] = X_{it}\beta + y_{i,t-1}X_{it}\alpha$.[1] The model resembles the conventional logit in that choice probabilities are modeled as the logit of covariate matrix X with effect parameter vector β. Additionally, transition models include all possible interactions with the lagged dependent variable. The interactions al-

---

[1]See also Jackman (2000, 8). Following a distinction introduced by Cox (1981), the model is an *observation-driven* transition model, as the realized choice $y_{i,t-1}$ is conditioned on. *Parameter-driven* transition models use the latent choice $y_{i,t-1}^*$ instead. While the latent model is often seen as preferable, it shall not be pursued here since estimation is notoriously difficult (Jackman, 2000, 9ff).

low the effect to vary dependent on previous choice. Accordingly, the αs express the change in β, given the previous choice. In practice, researchers deviate quite a bit from this general form due to how the number of interactions scales. Given the number of covariates, time points or choice alternatives, including all possible interactions makes the transition model very quickly too unwieldy to interpret. To arrive at a parsimonious model specification, insignificant or theoretically implausible interactions are typically eliminated (Diggle et al., 1994; Epstein et al., 2006).

While transition models for binary choices are most common, the approach travels easily to the ordered and multinomial case. These are merely more complicated in that an additional dimension is introduced. However, it is not quite clear to me how alternative-specific variables as in McFaddens conditional logit model, such as spatial distances, can be incorporated into transition models. Therefore, I turn to closely related econometric models which generally go by the name dynamic discrete choice models. These models were originally developed in marketing research to study consumer demand such as brand choice with scanner data of repeated product purchases (see Keane, 2013, for an overview). Formulated in the random utility framework, these models typically include product attributes such as price. Dynamics are introduced by allowing for state dependence. State dependence is "brand loyalty", or in the language of politics, party loyalty. The basic utility function of individual i to choose alternative j at time t may be written as

$$U_{ijt} = \alpha_j + X_{ijt}\beta + \omega_j y_{ij,t-1} + \epsilon_{ijt},$$

where α_j are alternative-specific intercepts, X_i is a matrix of choice attributes with utility weight parameter β. $y_{ij,t-1}$ is an indicator variable that takes the value of 1 if i chose j in the previous election, and 0 otherwise. Note that the parameter ω varies over j, which allows for different state dependence given different previous vote choices. This captures the notion that it matters which particular party is the party of origin. If state dependence has a positive effect, individuals obtain additional utility from choosing the same alternative again. If ω is negative, choosing the same alternative yields lower utility and individuals are more inclined to switch to a different alternative as these relatively gain utility. However, it is well established that such a "lagged dependent variable" specification will not only capture "true" state dependence, the *causal* effect of previous choice on present choice, but also unobserved or unspecified factors that affect both the vote choice in the European election and vote choice in the past national election, i.e., unobserved heterogeneity. This has been referred to as 'spurious' state dependence (Heckman, 1981). While a multitude of sophisticated statistical models have been developed to disentangle true from spurious state dependence (see e.g. Erdem, 1996), this paper shall remain agnostic to this question. That the lagged dependent variable will ultimately capture spurious state dependence is of no primary concern for my analysis. Where state dependence comes from is unimportant insofar as only its presence is relevant to explain which voters stay loyal or not. Putting it the other way around, effectively "controlling" for unobserved time-constant factors that co-

determine vote choice in both elections is a positive property of this specification. Unobserved variables "are not relegated to the disturbance term" (Jackman, 2000, 9), but are incorporated into the state dependence term, making the specification more robust to omitted variable bias.

Assuming the i.i.d. error-term ϵ to be distributed Type-1 extreme value, the model can be estimated as a multinomial logit for alternative-specific variables, or conditional logit model (McFadden, 1974). Collecting the terms of the utility functions in V_{ij}, the choice probability of voter i for party j is given by $Pr(Y_{ij} = 1) = \frac{e^{V_{ij}}}{\sum_{j=1}^{J} e^{V_{ij}}}$. Parameter estimates are obtained by maximizing log-Likelihood w.r.t. the parameters α_j, β and ω_j.[2]

How is the proposed model related to transition probabilities, and therefore vote switching? Imagine a very basic model that includes only party-intercepts α_j and state-dependence terms $\omega y_{ij,t-1}$. We can obtain each cell entry of the sample transition matrix by inserting the estimated coefficients into the multinomial logit function. $p_{A_2}|A_1$, a vote for party A given a vote for party A in the previous election is $e^{\alpha_1 + \omega_j * 1} / \sum_{j=1}^{J} e^{\alpha_j + \omega_j * y_{j,t-1}}$, etc. In a next step, transition probabilities are made dependent on other variables as well. In line with the spatial voting models so far used in this dissertation, I specify $X_{ijt}\beta$ as the squared Euclidean policy distance on the EU integration and Left-Right dimensions: $(p_{j_{eu}} - v_{i_{eu}})^2$ and $(p_{j_{lr}} - v_{i_{lr}})^2$. The respective effect parameters, the dimensional weights, are β_{eu} and β_{lr}. Note that, in line with spatial theory, β_{eu} and β_{lr} are modeled as parameters that are homogeneous in the population, meaning that every voter uses the same parameter to evaluate every party on EU integration issues, and the same parameter to evaluate them on left-right issues. To keep the model parsimonious, I also assume for the moment that the effect of spatial weights on transition probabilities does not depend on previous vote choice, all voters use the same spatial coefficients, irrespective of their party of origin. The sensitivity of the results to this assumption are tested in a second step. Additionally, I control for the partisan identity of the voters, as partisan voters should be suspected to exhibit a much higher state dependence than non-partisans. I include a lagged measure for party identification, i.e., party identification as it was reported by the respondents at the point of the Federal election. I include party identification as a choice-specific covariate that takes the value of 1 if voter i identifies with party j, and 0 otherwise.

7.4 Results

Table 7.3 reports parameter estimates and standard errors for three specifications of the dynamic discrete choice model of vote choice in the EP election. The first

---

[2]I estimate the models using the `mlogit` R package (Croissant, 2012; Henningsen and Toomet, 2010)

	EP Vote | Federal vote		
	LR only	LR & EU	Lagged Preferences
Distance EU^2		-0.045^* (0.005)	-0.030^* (0.004)
Distance LR^2	-0.069^* (0.007)	-0.068^* (0.007)	-0.045^* (0.006)
PID$_{t-1}$	0.937^* (0.105)	0.932^* (0.107)	0.940^* (0.105)
$Y_{t-1} = 1$	1.570^* (0.163)	1.572^* (0.164)	1.563^* (0.162)
$Y_{t-1} =$AfD	1.033^* (0.315)	0.730^* (0.336)	0.685^* (0.321)
$Y_{t-1} =$FDP	0.300 (0.386)	0.297 (0.386)	0.323 (0.384)
$Y_{t-1} =$Free Voters	0.605 (0.361)	0.587 (0.362)	0.605 (0.360)
$Y_{t-1} =$Greens	0.182 (0.289)	0.018 (0.293)	0.171 (0.291)
$Y_{t-1} =$Left	0.434 (0.392)	0.383 (0.394)	0.181 (0.383)
$Y_{t-1} =$Pirates	1.961^* (0.536)	2.221^* (0.538)	2.057^* (0.536)
$Y_{t-1} =$SPD	-0.585^* (0.249)	-0.556^* (0.252)	-0.554^* (0.249)
Observations	1,445	1,445	1,445
Log Likelihood	-1,656.827	-1,603.517	-1,664.832

Note: CSU is reference category. *p$<$0.05. Intercepts included, but not reported

Table 7.3: Dynamic discrete choice model estimates

model only includes squared distances on the Left-Right as an independent variable. Transition probabilities are also made dependent on squared EU integration distances in the second model. In the third, distance measures are replaced with distance measures calculated using lagged preferences, i.e., the EU and LR self placements reported by the respondents at the time of the Federal election. Analogous to my bivariate analysis, lagged preferences help exclude the possibility that the association between EU integration preferences and transition probabilities merely reflects voters changing their positions between elections.

The coefficients may be interpreted as follows. The coefficient for Y_{t-1} captures the state dependence for the reference party, the CSU. The coefficient is strongly positive. This was to be expected since the lagged choice indicator does not only capture the causal effect of the previous vote in the sense of a habit, but "sucks up" the effect of all unobserved or unspecified factors related to CSU vote choice in the Federal and CSU vote choice in the EP election. The remaining party-specific state dependence coefficients have to be interpreted relative to the coefficient for Y_{t-1}. One can therefore infer from the model estimates that former AfD and Pirate Party voters exhibit a statistically significantly higher state dependence, while former SPD voters have a lower state dependence than CSU voters.[3] The PID_{t-1} coefficient can be best understood as a modification to the utility to vote for the party a voter identifies with. As expected, party identification has a strong positive effect throughout the models. Two situations may be distinguished. If the party a voter identifies with is also the party chosen in the Federal election, the PID coefficient may be counted as an addition to the state dependence term, making the voter more loyal. However, if the party a voter identifies with is not the party chosen in the Federal election, it becomes more likely that the voter switches to the party he identifies with in the EP election. The dimensional weights, the coefficients of the squared distance terms, express the influence of spatial distance considerations on the probability of transitioning to one party, given the party chosen in the Federal election. The negative signs of the estimated weights indicate that increases in the distance between a voter and a particular party on both dimensions decreases the probability of transitioning to this party, keeping the distances to all other parties constant.

Comparing the first and second model specification, the magnitude of the coefficient for distances on the Left-Right dimension remains virtually identical. The coefficient for EU distances is statistically significant in the second model, indicating that voters factor in EU integration preferences when voting in EP elections. Analogous to the previous chapters, the relative dimensional weight can be calculated. For model 2 the (average) relative weight of EU integration preferences in the spatial utility calculus is estimated at .4, with a 95% confidence interval

---

[3]The large state dependence for Pirate Party voters might seem dubious at first sight, given the low retention rates that are reported in Table 7.2. However, the high state dependence needs to be interpreted relative to the Pirates low intercept estimate (-2.4), which then generates the accurate retention rate.

of [.33;.47]. This means, if one is comfortable with assuming that the distance measures are on the same scale, that EU integration preferences make up between a third and a half of the spatial considerations as they relate to transition probabilities. Spatial considerations related to European integration seem to play a noticeable role in determining voting transitions. However, they do not dominate over Left-Right considerations. This finding is very much in line with the findings for the general vote choice function presented in Chapter 5. The third model, that tests the sensitivity of the findings to temporal changes in voter preferences, come to a similar conclusion. Both distances between party positions and lagged self placements on EU integration and Left-Right are found to have a statistically significant negative impact on transition probabilities. While the magnitude of the coefficients is slightly lower, arguably due to the increased measurement error due to using lagged preferences, the estimated relative weight is almost the same: .4 [.3;.5]. Model 3 can therefore be counted as empirical support that the association between EU distances and transition probabilities is not due to voters having changed their views on European issues between elections.

Robustness Testing

The analyses so far conducted has been based on a sparse model specification that built on the assumption that the effect of policy distances is independent of the party chosen in the Federal election. While this assumption seems plausible from the view point of spatial theory, the assumption may obfuscate important patterns in the data that might invalidate my inference. One way of getting rid of the assumption is to incorporate additional interaction terms into the model. A simpler alternative approach is to estimate separate models for each set of respondents that voted for one specific party at the Federal election. While a joint model is certainly preferable in most instances, for my analysis which is conducted in the spirit of robustness testing, the latter shall suffice. Table 7.4 presents the obtained spatial estimates of the eight separate models.

Overall, the findings from the separate models suggest that the homogeneity assumption about the spatial effects on transition probabilities is tenable. For all sets of respondents who voted for the same party in the Federal election, the effect of both distances on the Left-Right and EU integration has a negative sign. Moreover, the spatial weights are confidently distinguishable from zero for most parties of origin. Since comparing coefficient magnitudes across the separate models may be reliable, I also calculate the relative weight of EU integration preferences, which should be more robust. Comparing the relative weights across the separate models, one can see that these vary between the models, but are consistently estimated as between roughly one third and three quarters, for the parties of origin with a sufficient number of observations to make authoritative inferences. In contrast, the joint model estimate (.4 [.3;.5]) was slightly more conservative about the weight of EU integration distances relative to the weight of Left-Right distances. Taking

	EP Vote | Federal vote			
	CSU	SPD	Greens	AfD
Distance EU^2	-0.052^*	-0.086^*	-0.090^*	-0.075^*
	(0.010)	(0.015)	(0.036)	(0.030)
Distance LR^2	-0.052^*	-0.040^*	-0.033	-0.035^*
	(0.009)	(0.009)	(0.017)	(0.012)
Relative Weight	0.49	0.68	0.73	0.69
	[0.36; 0.61]	[0.55; 0.79]	[0.38; 1]	[0.35; 0.88]
Observations	652	322	126	119

	FDP	Left	Free Voters	Pirates
Distance EU^2	-0.073	-0.074^*	-0.055	-0.045
	(0.043)	(0.022)	(0.042)	(0.042)
Distance LR^2	-0.091^*	-0.035^*	-0.034	-0.064
	(0.026)	(0.016)	(0.029)	(0.034)
Relative weight	0.45	0.67	0.61	0.43
	[0.0; 0.69]	[0.45; 0.93]	[0.0; 1]	[0.0; 1]
Observations	82	64	53	27

Note: *p<0.05. Intercepts and PID included, but not reported.

Table 7.4: Robustness Test: Spatial weight estimates from separate models for subset of respondents with same party of origin.

all this in the account, I think the separate model estimates have predominantly shown that spatial weights are substantively similar across parties of origin, rather than having shown that there is important variation in spatial weights. From this perspective, the potential gain in predictive accuracy that could be realized by allowing spatial weights to vary according to the party of origin is outweighed by the analytical value of a more parsimonious description of the data involving a homogeneous spatial weight. Accordingly, I choose to stick with the joint *LR & EU* model in the following.

EU preferences and vote switching: A simulation study

My analysis has so far established that EU integration preferences of voters influence transition probabilities. However, it has not become clear how these preferences are related to the phenomenon of vote switching. How strongly, and under what circumstances does policy disagreement on EU integration preferences lead to vote switching in EP elections? The model estimates can tell us only very little about that. This is because the interpretation of the spatial coefficient in terms of "a one unit change in distance leads to..." does not reflect how voting works in the real world. From the perspective of a real-world voter, distances can not be varied, e.g., in the sense that the voter moves one party's position to the left by one unit. For the voter, party positions are simply given. Alternatively, if the policy positions of the parties are given, and the voter's ideal point changes by one unit, obviously this affects the distance not only to one party, but the distance to all the other parties as well. Therefore, to determine the consequences of "moving the voter's ideal point" for vote switching, or comparing the likelihood of two voters to vote switch whose ideal points differ, requires an evaluation of the utility differential between all available parties of destination. This can not be inferred from the regression table.

Effectively this means that a specific choice situation needs to be considered in order to gauge the effect of distances on the EU integration dimension on vote switching behavior. To calculate the probability of a vote switch based on my model estimates, I need to know the party a voter voted for in the Federal election, the voter's ideal point on both dimensions, and the position of all available parties on both dimensions, and whether the voter has a party identification. Furthermore, the estimation uncertainty I have about the estimated parameters, that relate the above mentioned factors to the utility calculation of voters, needs to be incorporated. Simulation techniques (King et al., 2000) are an ideal way to accomplish all of this. Given a specific choice situation, the probability of transitions to all available parties of destination is calculated based on the estimates for the *LR & EU* model. If the party of destination is the party that was chosen in the previous election, one minus that transition probability is the probability of switching. By dividing the transition probability of the parties that are not the party of origin by the switching probability, the probability of the destination
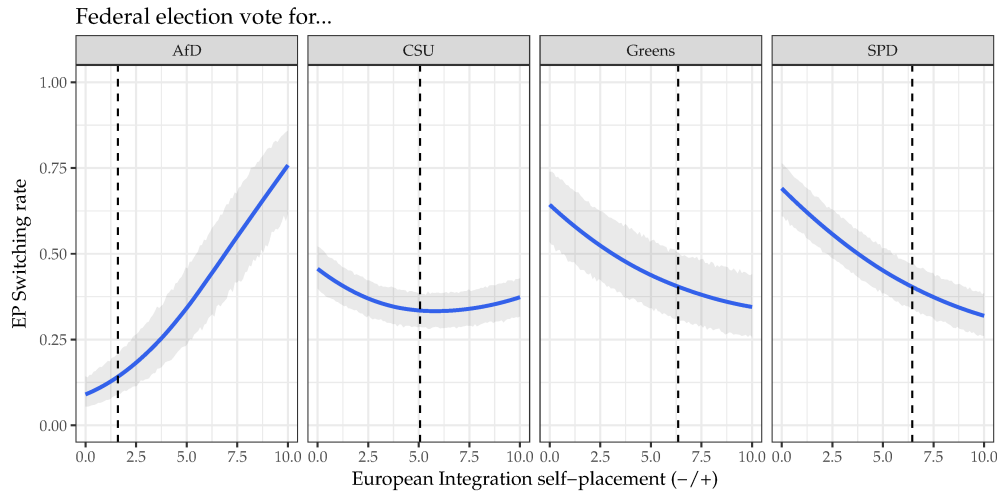
Figure 7.5: Simulated Model Predictions: EP switching rates

given that a switch takes place is obtained. Estimation uncertainty is incorporated by repeating this procedure for a large number of random draws from the joint sampling distribution of parameter estimates.

In a first simulation, I investigate the switching probabilities for the four main parties. Basically, the simulation setup replicates the idea behind Figure 7.1. I generate four sets of hypothetical voters that vary in their party of origin - AfD, CSU, Greens and SPD. Within each set, the hypothetical voters are made to vary in their EU integration ideal points across the whole spectrum from 0 to 11. Unlike in the descriptive bivariate analysis, I am now able to keep other factors constant. I set the Left-Right ideal point of all voters within each set to the position of the respective party. This means that all hypothetical voters are perfectly aligned with their former party on Left-Right issues - they have a distance of zero to their party of origin on the Left-Right dimension - and only differ in their distances on the EU integration dimension. Moreover, all hypothetical voters are assumed to have no party identification for any party. I use the party positions as reported in Chapter 3.

Figure 7.5 displays the switching probabilities obtained in this simulation. First of all, it is to be remarked that the patterns that arise from the stylized model and simulation closely resemble the patterns that I uncovered in the descriptive analysis of the observed data (see Figure 7.1). The probability to switch is predicted to rise monotonically with increasingly pro-EU positions for former AfD voters, and drops monotonically for former Greens and SPD voters. The U-shaped relation between EU integration preferences and switching for former CSU voters is also recovered by the simulation. Extreme anti-EU CSU voters are predicted to be more likely to vote switch in EP elections than extreme pro-EU voters, a pattern that was also observed in the bivariate analysis of the raw data. The findings of the first simulation are encouraging. The substantive features of patterns observed
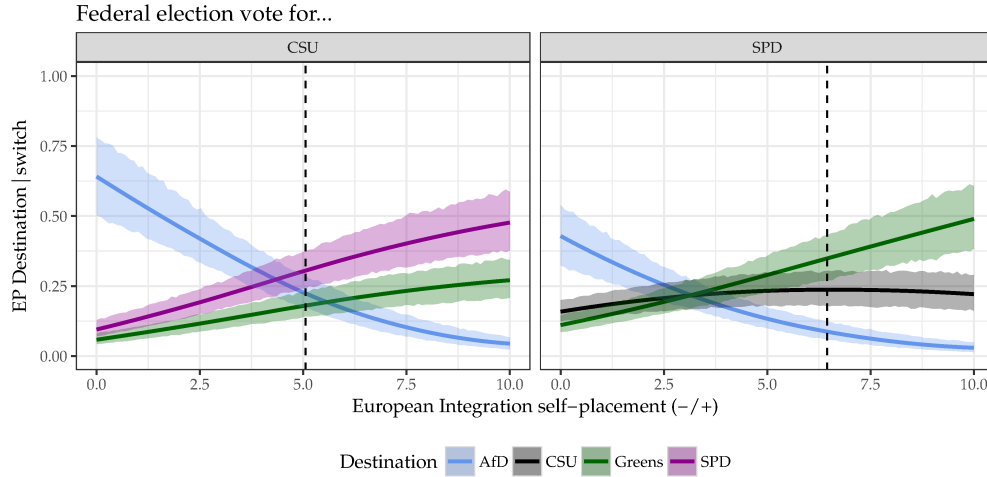
Figure 7.6: Simulated Model Predictions: Probability of EP destination given a switch.

in the raw vote switching data can be surprisingly accurately recovered using a very parsimonious model of multidimensional spatial voting. In order to scrutinize my model further, I conduct a second simulation study which investigates the predicted destination for switchers. The simulation set up remains unchanged, however I now calculate the predicted probability of destination given that a switch has taken place, as described in the above. These are plotted in Figure 7.6.

For about two thirds of anti-EU voters switching away from the CSU, the predicted destination is the AfD. The AfD is not a viable destination for pro-EU CSU switchers. Here around one half are predicted to switch to the SPD, and around one quarter to the Greens, both of which hold the most pro-EU positions in the party set. Anti-EU voters switching away from the SPD are predicted to be most likely to choose the rightist AfD, but less so than CSU voters, since former SPD voters have to trade away closeness on the Left-Right dimension to find their better match on EU integration. For pro-EU SPD voters, the Greens are predicted to be the main destination for the few vote switchers. The almost flat destination probability for the CSU is remarkable. On EU integration, the CSU at first sight appears as an attractive destination for SPD voters that are slightly more Euroskeptic than their parties. However, since voters put substantial weight on Left-Right distance, the benefit of voting CSU due to added closeness on EU integration preferences is completely canceled out by the fact that SPD voters would have to give up too much in terms of closeness on the Left-Right dimension to get that better match on EU integration.

The model predictions about the destinations of former CSU and SPD voters are also very much in line with the patterns observed in Table 7.2. Compared to the raw data, the minimum destination probabilities for all parties and the SPD probability for extreme pro-EU CSU voters seem slightly underestimated, the AfD

probability for extreme anti-EU voters slightly overestimated, which might hint at calibration issues.[4] However, there is a clear substantive similarity between the patterns in the raw data and the model predictions, which suggests that the model captures important features of the vote choice dynamics between the Federal and the EP election that are not explained by alternative theoretical explanations. Moreover, it is to be kept in mind that the model is extremely parsimonious. Of course there are quite a number of parameters - J-1 intercepts, the PID effect and J state dependence parameters, but these merely calibrate the model. The patterns in the association between the quantities of interest and spatial preferences are solely governed by two parameters that express the spatial weight of policy distances, and not the result of any bells and whistles.

7.5 Conclusion

This chapter of the dissertation has presented a study of the impact of policy preferences on European integration on vote switching between the Federal and EP elections in the German state of Bavaria. The main finding is that, contrary to the classical second-order interpretation of voting behavior, preferences specific to the context of the second-order election, namely preferences about European Integration, seem to play a substantial role in explaining patterns of vote switching between the Federal and the EP election. The patterns I find are consistent with a model of multi-dimensional spatial voting in which voter adjust the saliency of arena-specific policy dimensions. Moreover, the analysis in this chapter suggests that, in line with the theoretical model, whether EU integration preferences increase the likelihood of vote switching is fundamentally contingent on the existence of viable party alternatives on the dominant first-order dimension of political conflict, the Left-Right dimension. For example, while euroskeptic rightist former CSU voters are more likely to switch to the euroskeptic rightist AfD in EP elections, rightist CSU voters who are more pro-European than their party are less likely to switch because the pro-European parties SPD and Greens are leftist parties. The likelihood of switching intimately depends on the constellation of parties in the political space, relative to voter ideal points. In consequence, this means that vote switching due to EU preferences might play out very differently not only in different party systems, but also for different voters within the same electorate.

This finding calls into question the way in which vote switching has been studied so far. Research designs that rely on coarse categorizations of voting behavior and do not account for the characteristics of the parties that stand for election run the risk of missing crucial parts of the story. The research design and model presented in this chapter may help future research to surmount many of the difficulties of

---

[4]These are to be expected, as one should keep in mind that the simulation controls for Left-Right position and party identification. Of course this is not the case in the bivariate investigation.

studying voting behavior in consecutive elections across different electoral arenas. The simple transition model used in this paper can be easily extended to include other aspects of dynamic voting, and to rule out alternative explanations of why some voters vote for different parties in different elections. While my model has accounted for second-order factors by party-specific intercepts and state dependence, these factors may also be modeled explicitly. Most promising seems extending the model to include non-voting, which has so far been studied separately from switching.

Chapter 8

Conclusion

Modern democracies are almost exclusively polities that are organized in multiple levels of government. In multi-level democracies, citizens are not only called to the polls in presidential or national legislative elections, but also at the local, regional, subnational or transnational level. Following normative democratic theory, separate elections are necessary to legitimize policy-making at each specific level and to hold the elected accountable. Importantly, separate, level-specific legitimation is needed because policy-making at each level is concerned with different political issues. Separate elections allow citizens to express their preferences on precisely the issues that are at stake at a specific level: While a voter trusts one party on national-level issues, she may choose a different party at local elections if it better represents her views on local issues. However, whether voters act in this way is theoretically and empirically contested in political science research. While on the one hand it is cited as the reason why institutions are designed as they are and seen as a necessary prerequisite for true democratic legitimation, political science is highly skeptical whether voters are able and/or willing to distinguish between the levels of government on the other. So far, political science research has not satisfactorily addressed this discrepancy theoretically and has failed to develop convincing empirical strategies to arrive at more definite empirical answers to this important question: Do citizens actually make use of the possibilities that multi-level electoral democracy offers, and vote based on the specific political issues that are at stake at elections at the different levels of government?

The fundamental question whether voters pursue a "one-size-fits-all" approach to voting or whether they adjust their voting behavior calculus because they perceive that different issues are at stake at different levels has come to a head in research on European Parliamentary (EP) elections. It is here that the question whether voters differentiate between the different levels of government has culminated in a long-lasting academic debate about the role of genuinely European issues in EP elections. The traditional theoretical paradigm that has guided the analysis of electoral outcomes in EP elections, and by extension also dominates the study of multi-level electoral outcomes in general, the second-order elections

model, portrays elections at all other levels as subordinate to the national level. These second-order elections are conceptualized as following a nationalized logic that crowds out all factors that are specific to the second-order arena, such as the specific political issues or candidates at stake. Accordingly, EP elections are not about European issues, but about national-level issues. Second-order theory therefore by design sees multi-level democracy as fundamentally defunct. If it is indeed the case that citizens are unable to distinguish between the different levels of government, multi-level elections cannot fulfill their function of legitimizing policy-making or even hold the elected accountable.

Academic opposition to this pessimistic interpretation of how voters behave in multi-level elections has formed in the recent decades. Collected under the umbrella term "Europe Matters argument", a diverse set of contributions has argued, and presented empirical evidence that EP elections are not only less important national elections as stipulated by the second-order elections model, but have over the years become "more European". The European character of EP elections means that European citizens increasingly integrate genuinely European issues, such as their preferences over EU policy and their attitudes towards the European project into their decision making process at EP elections. Importantly, the literature implies that citizens do not only integrate European aspects in their overall "one-size-fits-all" voting decision calculus, but that these aspects play a heightened role *only* in EP elections. This would mean that voters possess a level-specific voting decision calculus and adjust their decision making calculus to account for the specific issues that are at stake in EP elections. In the grand scheme of things, this would mean that there is a democratic benefit to holding EP elections, as it allows EU citizens to express their level-specific preferences and to hold the politicians accountable that are elected at the European level.

The Europe Matters literature has laid important ground-work in working out the flaws and limitations of the second-order elections model, yet has so far fallen short of proposing a convincing alternative theoretical framework that is grounded in a general theory of voting behavior. To be fair, this is not a shortcoming of the literature on European elections, but of the political science discipline as a whole. As it stands, political science lacks an analytical model to analyze individual-level policy voting behavior across multiple or multi-level elections that would enable researchers to study the determinants of changes in individual voting behavior across elections or electoral levels. This is where this dissertation sets in. In Chapter 2, I propose a generalization of the multidimensional model of spatial competition to multiple elections that take place at different points in time and/or on different levels of government. The model distinguishes two mechanisms for policy-based change in voting behavior across elections: First, voter ideal points and party positions may simply change between elections, and lead to different electoral outcomes. Secondly, the weights that voters assign to the multiple relevant issue dimensions may vary across elections. Variation in weights encapsulates the idea that voters might care more about specific policy dimensions at specific

levels of government, which will incentivize them to change their voting behavior across elections. Following the second mechanism, electoral change is not a consequence of change in ideal points and party positions, but the consequence of a recalibration of the policy space across electoral levels. I employ this general model to study voting behavior in a typical multi-level elections setting, consisting of general national, subnational and EP elections. I place a particular focus on the role of European integration preferences in explaining electoral behavior in EP elections, and change in electoral behavior between subnational and national elections and EP elections. The relevant European policy space is defined as two-dimensional, with voter preferences being structured by an EU integration dimension and a Left-Right dimension. Based on this model, the conditions under which changes in relative dimensional weights impact electoral outcomes in multi-level elections are worked out. Three types of electoral outcomes are examined: Differential electoral participation, level-specific party vote choice and vote switching between elections. Accordingly, three substantive research hypotheses are investigated: Firstly, that voter preferences on EU integration can explain differential turnout in EP elections, secondly, that voters care more about EU integration preferences when choosing parties in EP elections than in national or subnational elections, and thirdly, that the views EU integration preferences motivate vote switching between national and EP elections. Additionally, the dissertation reflects on a critical assumption that governs the relationship between preference dimensions in the multidimensional model of spatial competition, the assumption of non-separable preferences. Accordingly, a fourth research question that this dissertation seeks to answer is whether Left-Right and EU integration preferences are separable.

Before empirical investigations into these four research questions are launched, Chapter 3 takes a step back and reflects on the fundamental inferential problems that empirical research on multi-election voting behavior faces. I clarify the identifying assumptions researchers need to make, why panel data is categorically needed and what the particular task is that panel data can fulfill in coming closer to reliably answering my research questions. At the core of my reflexions is the elimination of potential confounding due to unobserved heterogeneity. An overall research design that guides the empirical investigations of this dissertation is developed, which allows to eliminate heterogeneity bias even with the data limitations that multi-election voting behavior research is usually confronted with. I find that the proposed conditional likelihood approach to fixed effects models for discrete data is the most promising avenue for empirical research in the foreseeable future. Advantages and shortcomings of the conditional likelihood framework are discussed. Finally, the specific empirical case and panel data set that this dissertation employs is introduced.

Chapter 4 analyzes whether the EU integration preferences of voters can explain differential electoral participation in EP elections. It is the first study that seeks to uncover the causal mechanism that connects EU integration preferences

and electoral participation in EP election with panel data. A research design is developed that leverages available panel data to the fullest to overcome some of the gravest inferential problems that studies of individual-level turnout on observational data face. My empirical findings suggest that pro-EU integration preferences make electoral participation in EP elections more likely in the empirical case of investigation, but only for voters who locate themselves left of the center on the traditional dimension of party conflict - the Left-Right dimension. The relationship is non-existent for voters located to the right of center, and might even be negative for right-wing voters. The established conditionality of the relationship on the Left-Right placement of voters is an important finding that gives a valuable insight into the mechanism that links EU preferences and EP election turnout. While the finding is clearly not consistent with conventional theoretical explanations put forward by the second-order elections model or the legitimacy argument, it meets the predictions derived from a multidimensional model of spatial competition that explains differential abstention in EP elections by increased policy alienation arising from the recalibration of the policy space.

Chapter 5 analyzes the role of EU integration preferences in the voters' party choice calculus at multi-level elections. The diverging hypotheses brought forward by the second-order elections model and the Europe Matters argument are precisely formulated in the multidimensional model of spatial competition. To do so, I introduce the novel theoretical concept of relative dimensional weight. The second-order model's transfer hypothesis predicts a constant relative weight of EU integration preferences across the electoral levels, while the Europe Matters argument's recalibration hypothesis predicts the relative weight to be higher in EP elections than in national or subnational election. A research design is developed that allows for the consistent estimation of relative dimensional weights from individual-level panel survey data. The empirical results indicate that there is systematic variation in the spatial voting calculus of Bavarian voters between the elections. Although there is substantial uncertainty that is associated with the estimates, the evidence is more consistent with the recalibration hypothesis than the transfer hypothesis: Voters seem to put a larger emphasis on EU integration preferences when they cast their vote in EP elections than in other elections. This finding may be counted as evidence that voters are likely able and/or willing to distinguish the different levels of government, and to adjust their voting behavior in the level-specific elections accordingly. However, the evidence also indicates that the extent to which voters recalibrate is moderate at best. European issues do not play a dominant role in EP elections, as the traditional dimension of political conflict, Left-Right, maintains its role across multi-level elections.

Chapter 6 sheds light on a far-reaching and often neglected assumption on which the multidimensional model of spatial competition relies - that policy preferences on multiple dimensions are separable. The chapter consists of an article published in *Political Analysis* (Stoetzer and Zittlau, 2015) that works out the theoretical implications and statistical consequences of omitted non-separability, and pro-

poses a model to estimate non-separability. In the second part of the Chapter, I employ the developed methodology to test the robustness of the analysis conducted in Chapter 5 to potential non-separability. The analysis constitutes the first study that investigates whether Left-Right and EU integration preferences are separable in the spatial voting calculus of European voters. Moreover, the analysis implements the methodology developed in Stoetzer and Zittlau (2015) in the panel fixed-effects model of Chapter 5, and therefore constitutes the first study of non-separability with panel data. The empirical evidence suggests that policy preferences on Left-Right and EU integration are separable, lending additional credibility to the findings presented in Chapter 5.

Chapter 7 investigates the incidence of vote switching in EP elections as a consequence of a recalibration of the policy space. Complementing the fixed effect discrete choice approach used in Chapters 4, 5 and 6, this chapter proposes an alternative methodology to analyze changing voting behavior that centers on the modeling of transition probabilities. In a descriptive analysis of voting sequences in my panel data on national and EP election vote choice, I show that vote switching rates systematically covary with voter preferences on EU integration and distinctive party-specific vote switching patterns can be observed. In a next step, I translate my descriptive approach into a more formalized statistical model of multidimensional spatial voting. I find clear empirical evidence that EU integration preferences have a substantial impact on transition probabilities. Making use of simulation techniques to interpret my model results, I find that the vote switching patterns that my model predicts are very much consistent with the party-specific vote switching patterns observed in the data. This would suggest that some voters choose different parties in EP elections than in national elections because these better represent voter preferences on European integration issues, on which voters place a larger weight in EP elections than in national elections.

8.1   Summary and discussion of findings

Putting together the empirical evidence that was presented in the Chapters 4 to 7, a consistent picture of the role of EU -related preferences in Bavarian multilevel elections emerges. The EU integration preferences of citizens seem to play a distinctive role in explaining electoral participation and vote choice in the EP election, and change in individual voting behavior across national and EP elections. The general theoretical framework that consistently explains these empirical phenomena is a two-dimensional model of spatial competition in the European policy space. At the heart of the explanation is the idea that voters recalibrate their issue priorities between elections, which is expressed in the theoretical concept of changing relative dimensional weights. While Bavarian voters base their policy-based decision making calculus in national and subnational elections almost exclusively on considerations that fall onto the traditional dimension of political conflict, the

Left-Right dimension, they systematically integrate considerations related to EU integration issues into their decision making in EP elections. The consequences of this reweighing of the policy dimensions for voting behavior can be clearly defined by the model: Voters that are not aligned on EU integration issues with the parties they usually vote for in national and subnational elections will experience more policy alienation in the recalibrated policy space in EP elections - if there is no attractive alternative party that better reflects their recalibrated preferences. This makes certain ideological subsets of the Bavarian electorate more likely to abstain from participating in EP elections than in national elections. The non-alignment on EU integration with the parties voters prefer on Left-Right may also translate into a differential party vote choice in EP elections: When EU integration preferences gain importance in EP elections, some Bavarian voters are motivated to vote for alternative parties that better reflect their recalibrated preferences.

While my findings show that EU integration preferences clearly matter for Bavarian multi-level electoral behavior, they also indicate that the relationship is more complex than previously thought. I find no evidence that EU integration preferences have a direct and unconditional effect on voting behavior. In line with my theoretical reasoning it is rather how EU integration preferences work in conjunction with voters' ideological dispositions on the Left-Right dimension of political conflict that proves decisive. Throughout the dissertation, I find that Left-Right preferences remain the dominant guideline for the policy-based decision making of Bavarian voters, even in EP elections. Yes, EU integration preferences clearly matter when voters make up their mind in EP elections, but whether they translate into differential voting behavior such as level-specific abstention or vote switching is conditional on voters' Left-Right preferences. My findings not only highlight the continuous importance of Left-Right ideology, but also underline the importance of party system factors. The structure of party competition in the European policy space is a decisive contextual factor that governs if and how EU integration preferences translate into outcomes. Even in multi-party systems with a substantial number of parties such as the Bavarian one, voters' opportunities to express their preferences at the ballot box are limited by the party system. Party positions on Left-Right and EU integration are not randomly distributed in the policy space, but are arranged in distinctive patterns. In Bavaria, Euroskeptic parties are located at the fringes of the left-right spectrum, center-right parties hold moderately Euroskeptic positions, and center-left parties are pro-integrationist. This constellation of parties in the policy space creates opportunities for some voters to express their EU integration preferences at the ballot box, and limits those of others. This is because of the policy trade-offs that particular ideological subsets of the electorate face. For example, rightist voters with Euroskeptic preferences find ample opportunity to update their voting behavior in EP elections by casting a vote for the Euroskeptic far-right, as they have to trade away relatively little in term of Left-Right proximity to get their closest match on EU integration. Rightist voters with pro-integration preferences are less fortunate, as they would have to

trade away too much proximity on Left-Right in order to get a closer match on EU integration. These voters are likely to stick with their usual party - EU integration preferences are effectively rendered ineffective in this case -, or abstain in EP elections because they feel more alienated. The relationship between EU integration preferences and political outcomes is therefore conditional on the constellation of parties in the policy space. This implies that if the party system changes, over time or in a cross-country comparativist sense, so does the relationship between integration preferences and voting behavior.

In summary, my analysis of Bavarian multi-level elections suggests two important lessons for our understanding of the role of EU integration preferences in national and supranational elections: Firstly, EU integration preferences matter, and voters put a larger emphasis on them in EP elections than in national elections. Secondly, whether and how they translate into electoral outcomes is fundamentally contingent upon voter preferences on Left-Right, and the spatial configuration of party competition. But what are the broader implications of my findings? Two aspects of this question may be distinguished: First of all, how do my conclusions travel beyond the empirical case of investigation, to other EU member states and other time periods? And secondly, what can my conclusions tell us about the overall research question about the functioning of multi-level democracy in general?

As the research designs presented in this dissertation have purposely focused on guaranteeing the internal validity of inferences, generalizability was necessarily neglected. It might therefore be argued that my conclusions do not travel beyond the particular case under study, the Bavarian multi-elections 2013-2014. I think this is an extreme and unhelpful view. To arrive at a more nuanced evaluation, two aspects may be distinguished. Firstly, it is clear that the narrower conclusions, the estimated effect sizes and specific behavioral patterns, are unlikely to replicate precisely in other electoral contexts. Indeed my theoretical work pressingly indicates that this should not be the case: Variation in the party system should indeed change the relationship between EU integration preferences and electoral outcomes. EU integration preferences can be muted if parties form an "issue cartel" on EU integration, or activated by issue entrepreneur parties that exploit representational gaps in the party system. Moreover, it seems plausible that the degree of politicization of EU issues varies substantially between electoral contexts, and is influenced by event shocks and electoral campaigns. This should modulate the relative weight of EU integration preferences, and the magnitude of recalibration between national and EP elections. In the end it is an empirical question how heterogeneous effect sizes are across different electoral contexts, about which this dissertation needs to remain largely silent and which is to be left to future research to explore. The second aspect of the generalizability question concern the broader conclusions this dissertation draws concerning the existence and nature of the relationship between EU integration preferences and multi-level voting behavior and its underlying mechanism. I think the work presented in this dissertation provides ample reason to be comfortable with transferring the broader lessons learned to

other electoral contexts as well. Employing research designs that produce reliable inferences by effectively eliminating potential confounding, I find consistent evidence that EU integration preferences matter to Bavarian voters. It is therefore convincingly shown that voters have the ability to adjust their EP election voting behavior to the EP election-specific issues that are at stake, which motivates them to abstain and vote for different parties in EP elections than in national elections. If Bavarian voters can do it, why shouldn't voters in other EU member states? Given that I see no reason that is grounded in solid theory, I think it is implausible to suggest that the Bavarian electorate in the 2013-2014 multi-level elections is fundamentally different from other electorates or time-periods. Accordingly, the basic mechanism that relates EU integration preferences to European multi-level election behavior should exist in other electoral contexts as well. Nevertheless, it is clear that this dissertation is not the final word on this research hypothesis. Further empirical analyses of European multi-election voting behavior in other electoral contexts that pays close attention to the inferential challenges outlined in this dissertation is imperatively needed to increase confidence in the conclusions.

If one accepts the proposed conclusions about the EU integration-multi-level election nexus, what are the implications for the broader research question concerning the functioning of multi-level democracy? Does multi-level democracy only guarantee procedural democracy, or does it fulfill its promise of improving substantive democracy by enabling level-specific representation and holding the elected accountable at each separate level? This vital question hinges upon the empirical question whether real-world voters are able and willing to distinguish what is a stake at the different levels, and to act upon these level-specific evaluations. I think this dissertation has convincingly worked out that the pessimistic view on multi-level democracy, exemplified by the second-order elections model's transfer hypothesis, is not tenable on both theoretical and empirical grounds. The transfer hypothesis does not rest upon a convincing theoretical argument why level-specific considerations should be systematically crowded out in second-order elections. Empirically, voters indeed seem to be able to identify the specific issues that are at stake in EP elections, and adjust their decision making calculus accordingly. This dissertation therefore suggests that voters are able to fill multi-level electoral institutions with life by making use of the opportunities they provide. In this light, multi-level electoral institutions do not only guarantee procedural democracy, but are a necessary feature of modern substantive democracy.

While I do not see a convincing argument why my overall conclusion would not be applicable to other multi-level second-order elections as well, it is clear that this generalization necessarily remains uncertain. I am however confident: If voters are willing and able to act upon level-specific policy preferences in European elections, which are arguably cognitively challenging since EU-related issues often remain abstract and the consequences of policy decisions opaque to the average voter, why shouldn't they do so in other multi-level elections as well? If level-specific issues in other second-order elections, such as education policy in the German states,

are much closer to the average voter, wouldn't one expect even more level-specific voting behavior? While this seems plausible at first sight, one should be cautious to draw this conclusion. Even though the cognitive challenge to voters may be higher on the one hand, it may on the other hand be easier to detect these effects in the European arena, since level-specific policy preferences in the European arena cross-cut the traditional policy dimension of national-level political conflict. In other arenas, where party competition has to a large degree integrated level-specific issues into the Left-Right policy dimension, such as education policy at the subnational level, the observable consequences of level-specific voting might be less pronounced, and difficult to detect: Even if voters put a larger emphasis on level-specific issues in such a situation, it will lead to the same result as if voters were to base their vote only on Left-Right. Effects might therefore be small and difficult to detect in these instances, given the data limitations and weak identification strategies that empirical research in this field commonly faces. Failure to detect these effects in observational data may therefore not necessarily mean the absence of level-specific policy voting.

8.2 Broader Research Contribution

This dissertation not only informs the narrower field of research on EP and multi-level democracy, but is relevant to the wider political science community, particularly to fields that rely on a spatial model of politics in their theoretical work, be that in the field of Policy Analysis or International Relations. In particular, this dissertation speaks to all scholars that employ multidimensional spatial models and seek to better understand the relationship between individual policy dimensions and the consequences of changes in how dimensions are related over time or in different situations. One key contribution of this dissertation is its re-appreciation of the theoretical richness of the multidimensional model of spatial competition. It picks up important aspects such as the notion of relative dimensional weights and non-separability of preferences that used to be integral parts of the model as it was originally conceived in the '60s and '70s, but have since fallen through the cracks of the academic mainstream. The dissertation also offers a number of new ideas, which naturally integrate into the classical spatial model. First and foremost, the application of spatial theory to a multi- or multi-level election context necessarily requires a dynamic and longitudinal perspective on spatial voting. Surprisingly, such an analytical framework has so far not been developed. This dissertation is the first contribution to do so by extending the spatial voting model to multiple elections and multiple electoral arenas. As it turns out, the longitudinal multidimensional spatial voting model that is proposed offers a concise, yet versatile framework to explain individual electoral change between elections as a consequence of changing policy priorities. How policy priorities change can be conveniently captured by the important theoretical concept of relative weight of

the policy dimensions. While relative weights are important quantities of interest in their own right, an interesting and important implication of spatial theory is that changes in relative dimensional weights between elections or electoral arenas have different electoral consequences for different individuals. Whether a voter abstains from voting, votes for the same party as last time, or changes his vote to another party fundamentally depends on the individual's multidimensional preferences relative to the multiple party alternatives in the multidimensional policy space. When dimensional weights change between elections, multi-party competition can lead to relatively complex patterns of individual electoral behavior. This insight is certainly not novel, but tends to be under-appreciated in many applications of spatial theory in my opinion. All too often, academic contributions rely on oversimplified or even incomplete spatial models. To the contrary, I think that acknowledging complexity reminds us of one of the major strengths of spatial theory: Its ability to describe and explain seemingly complicated behavioral patterns with a parsimonious model.

Furthermore, the dissertation provides novel insights into bridging the gap between spatial theory and their statistical investigation, and updates the statistical models used to implement the theoretical models. One key idea is expressing the relative dimensional weights that is deeply encoded in the spatial framework by setting estimated coefficient parameters into relation to each other. As this dissertation demonstrates, the resulting statistic may be a quantity of interest for many research questions. Additionally, by analyzing relative weights one arrives at a more meaningful interpretation of coefficient values that goes beyond the very limited scope of the Null hypothesis testing framework that has traditionally driven the interpretation of empirical results of statistical spatial models. This allows for making more meaningful statements about the role that particular policy preference dimensions play when voters make up their mind whom to vote for in an election. The dissertation also presents joint work with Dr. Lukas Stötzer on a statistical framework to integrate the concept of non-separability of preferences into statistical applications of the multidimensional spatial model. The non-separability statistic constitutes an additional dimension of analysis that may be relevant for a variety of research questions. Our research therefore opens up additional interesting and relevant avenues for hypothesis development and testing. Finally, the dissertation introduces transition models into the spatial analysis of multidimensional spatial voting in multiple elections. Transition models offer a consistent framework to study individual-level behavioral change across multiple elections, and readily translate into a statistical model. Surprisingly, this connection has not been made by existing research into policy-induced vote switching so far, an omission that this dissertation corrects. Rather than analyzing binary categorizations of voting behavior defined as vote switching that are oftentimes fundamentally at odds with the theory of spatial voting in multi-party systems, transition models allow to analyze switching in terms of transition probabilities, a notion that follows naturally from the spatial model. This framework allows to

connect spatial models of electoral change to a vast state-of-the-art econometric literature of consumer choice models that has remained largely unexplored by political science research. As I show in Chapter 7, such models are readily adapted to study the determinants of vote switching. As I show, this empirical strategy can be augmented with simulation techniques that greatly facilitate interpretation of the complex spatial patterns of vote switching implied by a spatial multi-election model of voting in multidimensional policy spaces. Additionally, I explore new graphical options that enable comparisons of empirical and estimated patterns of vote switching in order to determine model fit and answer the substantive research questions. In line with the overall theme of this dissertation, the transition model research design constitutes an important step towards appreciating the non-triviality and richness of spatial theory to describe and explain electoral choice.

Finally, the dissertation offers a number of contributions of a more or less purely statistical nature. First of all, the dissertation demonstrates how to harness the analytical power of panel fixed effect models for categorical dependent variables, in particular the class of panel conditional logit models. While panel random effect models have become popularized over the last decade, the political science literature as a whole seems to be largely unaware of the fixed effect option and its advantages. Panel fixed effect models can improve the identification strategy since they effectively deal with unobserved heterogeneity. Unlike random effects models, unobserved heterogeneity might be related to the independent variables in fixed effect models. Especially in electoral behavior research, unobserved heterogeneity is a prime concern since unobserved characteristics are likely to explain a vast proportion of behavior. Additionally, we have relatively little established knowledge of the relation between unobserved and observed characteristics which are often also measured very imprecisely. Assuming unobserved and observed characteristics to be unrelated might be heroic in this instance, and how this relationship is specified can have a great deal of influence on the substantive findings, and invalidate the drawn inferences. Another more practical advantage of the conditional likelihood framework is that it requires only a few panel waves - theoretically two panel waves are sufficient. In electoral research, where most panels are short and random effects not feasible, panel fixed effect models are often the only way to address unobserved heterogeneity. As I demonstrate in this dissertation, panel conditional logit models are very flexible and can be tailored to the specific research question at hand. Models for binary, ordered, multinomial dependent variables are available. Coefficients can be specified as varying across panel waves to model changes in the data generating process. Moreover, panel conditional multinomial logit can be adapted to incorporate alternative-specific independent variable. This option was not explored previously, making this dissertation the first contribution to propose what might be called a "double conditional" logit model, a version of McFadden's conditional logit model for panel data. This model promises to become the new workhorse statistical model for spatial voting research conducted on panel data. Combined with the research on non-separability, this dissertation

therefore promises to push the methodology of empirical applications of the spatial voting model to the next level.

8.3 Implications for future research

Given the various contributions that this dissertation makes, a number of open questions and guidelines for future research projects can be deduced - related to the narrower field of EP election and multi-level election research, for empirical applications of spatial voting models, and for electoral behavior research in general.

Firstly, research projects that launch more nuanced and broader, comparative investigations into the relationship between EU integration preferences and multi-level elections are called for. This dissertation has estimated average effects. While an average effect is representative of the electorate as a whole, effect sizes are likely to vary a great deal across individuals. Since the consistent estimating of individual-specific effects will remain impossible with observational data, researchers might seek to identify specific socio-demographic subsets or latent classes of the electorate that are more flexible in adjusting their voting behavior across levels than other groups that follow a more "one-size-fits-all" approach to multi-level voting. Political sophistication or psychological traits seem likely candidates for explaining individual-level variation in effect sizes. Future research projects addressing this research question will however have to address unobserved heterogeneity concerns, as unobserved factors are likely to be strongly related to these variables. Moreover, I only analyzed multi-level electoral behavior in one specific electoral context. Generalization to other contexts is therefore necessarily uncertain, especially since my theoretical argument implies that multi-level election behavior is contingent upon a context-dependent factor, the party system. A natural extension of the dissertation would therefore be a cross-country study that exploits the variation of party systems across EU member states. By analyzing whether country-specific relationships between level-specific preferences and voting behavior are in line with model predictions, additional confidence in the theoretical framework proposed in this dissertation can be realized. The proposed panel survey extension to the European Elections Studies Voter Survey component, that may become available in the near future, will provide an excellent opportunity to implement said research design. Not only cross-country variation, but also variation over time is of interest. An interesting research question might be whether the extent of recalibration in EP elections has increased over time. However, it is not clear to me how a convincing research design can be implemented with the data that is available at the present date. Another important question concerns the timing of second-order elections in the national electoral cycle. The second-order election model predicts that the extent of second-order effects varies over the cycle, and so might the extent to which voters perceive level-specific policy considerations to be relevant, and adjust their voting calculus accordingly. It seems

plausible that the closer EP elections are held to national general elections, the more EP electoral behavior is contaminated by national-level considerations. The different points in the electoral cycle on which an EP election falls in the EU member states offer a unique opportunity to study these potential crowding-out effects. A future study could for example estimate the magnitude of recalibration with the here proposed methodology, and investigate whether effect sizes are conditional on the timing of the elections. However, such a study would have to disentangle cycle- and country-specific effects, which may prove difficult. Finally, the dissertation only investigated the role of one kind of level-specific preferences, EU integration preferences, in one specific combination of multi-level elections, subnational-national-European. Evidence on the role of other level-specific policy dimensions or issues in other combinations of elections is needed to arrive at a more general picture of the role of level-specific issues in multi-level elections. The research design proposed here can be brought to bear on this question with relative ease. The German federal system may provide potentially interesting settings. States are responsible for education policy, and given suitable data, it could be tested whether voters put a larger emphasis on education policy aspects in State elections than in national-level elections. The comparison of local-level with national-level voting behavior may provide additional leverage on the broader research question about multi-level democracy. Survey data on the local level is rare, but quasi-experimental approaches that exploit plausibly exogenous variation between localities may prove viable here as well.

The second set of implications for future research concerns the empirical application of multidimensional models of spatial competition in voting behavior research, and potentially in other fields of the political science discipline as well. Spatial theory has been criticized for treating preferences as separable and exogenous, which is not logically consistent with the wider rational choice framework upon which spatial theory is built (Milyo, 2000). The spatial methodology developed in this dissertation addresses these important points of criticism, and therefore holds important lessons for the future of spatial voting research that will held to make models more realistic and empirical findings more reliable. Non-separability is an interesting theoretical concept that may open up new perspectives on the multi-dimensional preference structures that guide real-world voting behavior. Future research projects may leverage non-separability to answer long-standing academic debates, such as whether policy-based voting in US presidential elections is primarily driven by economic or socio-cultural policy preferences, how these policy preferences are related to each other and how this relationship has evolved over time. It is also clear that non-separability should constitute a vital part of robustness testing in future empirical applications of spatial theory to improve the validity of the drawn inferences. Non-separability may also prove to be of great relevance outside of the narrower confines of the proximity voting model, e.g., for the issue voting literature. Secondly, I would suggest that empirical applications of spatial voting models need to pay much closer attention to the role of preferences in

a causal model of electoral behavior. Policy preferences do not arise from nothing, but are to a substantial proportion influenced by non-policy factors, such as an individuals socio-demographic characteristics, income or religious beliefs. Conventional attempts of statistical control are likely to fail in isolating the "true" effect of preferences, since researchers only have crude measurements of controls, and do not know the correct model specification. Moreover, important factors are likely to always remain unobserved, such as voters' cognitive ability, risk acceptance or psychological inclinations. Due to the limitations of this strategy, the effect of non-policy factors on voting is likely to be wrongly attributed to the effect of policy preferences, which means that estimated parameters of the effect of policy distances are likely to be considerable overestimated in most settings. This leads me to conclude that future spatial models need to be estimated with fixed-effect models on panel data to successfully eliminate potential confounding, and improve the credibility of empirical evidence derived from applications of spatial theory.

Finally, the implications of potential confounding due to unobserved heterogeneity are not only relevant for the spatial voting literature, but for the wider field of electoral behavior research. These implications are not of a "wonky", purely statistical nature, but raise deeper substantive issues that the political science discipline faces. While panel analysis techniques belong to the standard repertoire of mainstream academic research in economics, psychology and sociology, they remain relatively infrequent in the political science discipline as a whole, and the narrower field of electoral behavior research in particular. The chief reason why the field has been so slow to adopt panel analysis is certainly the scarcity of panel designs in traditional voter survey studies. However, advances in online survey methodology have drastically reduced surveying costs, which means that more and more panel studies are becoming available. In all too many instances, not only in the field of electoral research, but in the political science discipline as a whole, researchers finally have access to long called-for panel data, yet continue to conduct their empirical analysis of panel data using cross-sectional data analysis techniques. By doing so, they fall short of making the best use of the panel structure of the data, whose main advantage is that it allows for addressing heterogeneity bias. One reason for this is certainly that statistical education has largely focused on cross-sectional techniques, which are then reinforced by the incentive structure of the discipline. Another, more practical reason is that the panel data analysis techniques that are part of the econometrics and psychometrics mainstream are built for panel data with a large number of waves and linear dependent variables, and have only a limited scope of application for typical political science panel data sets. Therefore the discipline needs to develop and popularize panel data analysis techniques that are tailored to the particular structure of political science data, i.e., for short panels with limited dependent variables. The conditional likelihood framework has the potential to become an important part of the solution. With a small number of panel waves, conditional likelihood is often the only game in town in terms of panel data analysis techniques. By effectively dealing with unobserved

heterogeneity, conditional likelihood allows for more reliable estimates of the effect of time- or election-varying variables on political outcomes. As heterogeneity is allowed to be related to the explanatory variables in any way possible, all potential time-constant confounding factors are effectively eliminated, which also facilitates model specification enormously. This makes the conditional likelihood framework particularly powerful for all research questions that are at particularly high risk of confounding, such as studies that seek to isolate the effect of candidate, party, coalition or policy evaluations on vote choice or turnout. Contrary to prevalent beliefs, panel techniques also lend themselves to study the effect of time-constant factors as long as effect parameters are specified as time-varying. For instance, this is often the case in research that is interested in how changes in electoral context lead to changes in the data generating process. Potential applications are studies interested in estimating whether the effect of socio-demographic factors on electoral outcomes was stronger in one election than another, or how institutional changes between elections affect the vote choice mechanism. In these instances, the employment of conditional likelihood techniques should at least become a required step in robustness testing. Thereby they can help reduce spurious findings and reporting of overestimates, and help improve the hygiene of the discipline. However, the clear analytical advantages that can be realized by the conditional likelihood framework come at a cost: The conditional likelihood framework is by design inefficient, i.e., more conservative statistical findings that are associated with larger uncertainty are obtained. While I am aware that publication standards unfortunately disincentivize conservative findings, I do not think that this should absolve researchers of trying to address heterogeneity. Larger uncertainty is the necessary cost researchers have to pay, but this uncertainty more accurately reflects how comfortable researchers really should be with making their conclusions. Yes, conditional likelihood techniques make research and its publication harder, but they promise to produce more trustworthy results, of which the discipline is direly in need.

Bibliography

Abrevaya, J. (1997). The equivalence of two estimators of the fixed-effects logit model. *Economics Letters 55*(1), 41–43.

Adams, J. F., S. Merrill III, and B. Grofman (2005). *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors.* Cambridge: Cambridge University Press.

Aldrich, J. H. (1993). Rational Choice and Turnout. *American Journal of Political Science 37*(1), 246–278.

Aldrich, J. H. and R. D. McKelvey (1977). A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections. *The American Political Science Review 71*(1), 111–130.

Aldrich, J. H., J. M. Montgomery, and D. B. Sparks (2014). Polarization and Ideology: Partisan Sources of Low Dimensionality in Scaled Roll Call Analyses. *Political Analysis 22*(4), 435–456.

Allison, P. D. (1994). Using Panel Data to Estimate the Effects of Events. *Sociological Methods & Research 23*(2), 174–199.

Allison, P. D. (1999). Comparing Logit and Probit Coefficients Across Groups. *Sociological Methods & Research 28*(2), 186–208.

Allison, P. D. (2009). *Fixed Effects Regression Models.* Thousand Oaks, CA: SAGE Publications, Inc.

Alvarez, R. M. and J. Nagler (1998). When Politics and Models Collide: Estimating Models of Multiparty Elections. *American Journal of Political Science 42*(1), 55–96.

Alvarez, R. M. and J. Nagler (2004). Party system compactness: Measurement and consequences. *Political Analysis 12*(1), 46–62.

American National Election Studies (2009). American National Election Studies (ANES) Panel Study, 2008-2009. ICPSR29182-v1.

Andersen, E. B. (1970). Asymptotic properties of conditional maximum likelihood estimators. *Journal of the Royal Statistical Society. Series B 32*(2), 283–301.

Andersen, E. B. (1973). *Conditional Inference and Models for Measuring.* Copenhagen: Mentalhygiejnisk Forlag.

Anderson, T. W. and L. A. Goodman (1957). Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics 28*(1), 89–110.

Andreß, H.-J., K. Golsch, and A. W. Schmidt (2013). *Applied Panel Data Analysis for Economic and Social Surveys.* Berlin, Heidelberg: Springer.

Ansolabehere, S. and E. Hersh (2012). Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis 20*(4), 437–459.

Arzheimer, K. (2015). The AfD: Finally a Successful Right-Wing Populist Eurosceptic Party for Germany? *West European Politics 38*(3), 535–556.

Bakker, R., C. de Vries, E. Edwards, L. Hooghe, S. Jolly, G. Marks, J. Polk, J. Rovny, M. Steenbergen, and M. A. Vachudova (2015). Measuring party positions in Europe. *Party Politics 21*(1), 143–152.

Bakker, R., S. Jolly, and J. Polk (2012). Complexity in the European party space: Exploring dimensionality with experts. *European Union Politics 13*(2), 219–245.

Bartels, L. M. (2006). Priming and persuasion in presidential campaigns. In R. Johnston and H. E. Brady (Eds.), *Capturing campaign effects*, pp. 78–112. University of Michigan Press.

Benewick, R. J., A. H. Birch, J. G. Blumler, and A. Ewbank (1969). The Floating Voter and the Liberal View of Representation. *Political Studies 17*(2), 177–195.

Benoit, J.-P. and L. A. Kornhauser (1994). Social Choice in a Representative Democracy. *The American Political Science Review 88*(1), 185–192.

Benoit, K. and M. Laver (2006). *Party Policy in Modern Democracies.* London: Routledge.

Bernstein, R., A. Chadha, and R. Montjoy (2001). Overreporting Voting. *Public Opinion Quarterly 65*(1), 22–44.

Berry, W. D., M. Golder, and D. Milton (2012). Improving Tests of Theories Positing Interaction. *The Journal of Politics 74*(3), 653–671.

Blais, A. (2006). What Affects Voter Turnout? *Annual Review of Political Science 9*(1), 111–125.

Blais, A. (2010). Making Electoral Democracy Work. *Electoral Studies 29*(1), 169–170.

Blais, A., J.-F. Laslier, and K. van der Straeten (2016). *Voting Experiments.* Switzerland: Springer International Publishing AG.

Blank, H., V. Fischer, and E. Erdfelder (2003). Hindsight bias in political elections. *Memory 11*(4-5), 491–504.

Blondel, J., R. Sinnott, and P. Svensson (1997). Representation and voter participation. *European Journal of Political Research 32*(2), 243–272.

Blondel, J., R. Sinnott, and P. Svensson (1998). *People and Parliament in the European Union: Participation, Democracy, and Legitimacy.* Oxford: Oxford University Press.

Boomgaarden, H. G., D. Johann, and S. Kritzinger (2016). Voting at National versus European Elections: An Individual Level Test of the Second Order Paradigm for the 2014 European Parliament Elections. *Politics and Governance 4*(1), 130–144.

Boomgaarden, H. G., A. R. T. Schuck, M. Elenbaas, and C. H. de Vreese (2011). Mapping EU attitudes: Conceptual and empirical dimensions of Euroscepticism and EU support. *European Union Politics 12*(2), 241–266.

Boskin, M. J. and F. C. Nold (1975). A Markov Model of Turnover in Aid to Families with Dependent Children. *The Journal of Human Resources 10*(4), 467.

Brambor, T., W. R. Clark, and M. Golder (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis 14*(1), 63–82.

Brams, S. J., D. M. Kilgour, and W. S. Zwicker (1997). Voting on referenda: the separability problem and possible solutions. *Electoral Studies 16*(3), 359–377.

Brams, S. J., D. M. Kilgour, and W. S. Zwicker (1998). The paradox of multiple elections. *Social Choice and Welfare 15*(2), 211–236.

Braun, D., S. Hutter, and A. Kerscher (2016). What type of Europe? The salience of polity and policy issues in European Parliament elections. *European Union Politics 17*(4), 570–592.

Braun, D. and M. Tausendpfund (2014). The Impact of the Euro Crisis on Citizens' Support for the European Union. *Journal of European Integration 36*(3), 231–245.

Brüderl, J. (2010). Kausalanalyse mit Paneldaten. In C. Wolf and H. Best (Eds.), *Handbuch der sozialwissenschaftlichen Datenanalyse*, pp. 963–994. Wiesbaden: VS-Verlag.

Calvert, R. L. (1985). Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence. *American Journal of Political Science 29*(1), 69–95.

Campbell, A., P. Converse, W. Miller, and D. Stokes (1960). *The American Voter.* New York: Wiley.

Caramani, D. (1996). The nationalisation of electoral politics: A conceptual reconstruction and review of the literature. *West European Politics 19*(2), 205–224.

Caramani, D. (2004). *The nationalization of politics: the formation of national electorates and party systems in Western Europe.* Cambridge: Cambridge University Press.

Caramani, D. (2011). The Europeanization of electoral politics: An analysis of converging voting distributions in 30 European party systems, 1970-2008. *Party Politics 18*(6), 803–823.

Carmines, E. G. and J. A. Stimson (1989). *Issue Evolution: Race and the Transformation of American Politics.* Ithaca, NY: Princeton University Press.

Carrubba, C. and R. J. Timpone (2005). Explaining Vote Switching Across First- and Second-Order Elections: Evidence From Europe. *Comparative Political Studies 38*(3), 260–281.

Cassel, C. A. (2003). Overreporting And Electoral Participation Research. *American Politics Research 31*(1), 81–92.

Chamberlain, G. (1980). Analysis of Covariance Data with Qualitative Data. *The Review of Economic Studies, 47*(1), 225–238.

Clark, N. and R. Rohrschneider (2009). Second-Order Elections versus First-Order Thinking: How Voters Perceive the Representation Process in a Multi-Layered System of Governance. *Journal of European Integration 31*(5), 645–664.

Clark, T. S. and D. A. Linzer (2015). Should I Use Fixed or Random Effects? *Political Science Research and Methods 3*(2), 399–408.

Clausen, A. R. (1968). Response Validity: Vote Report. *Public Opinion Quarterly 32*(4), 588.

Clinton, J., S. Jackman, and D. Rivers (2004). The Statistical Analysis of Roll Call Data. *American Political Science Review 98*(2), 1–16.

Coleman, J. S. (1964). *Models of change and response uncertainty.* Englewood Cliffs, NJ: Prentice-Hall.

Commission of the European Communities (2012). *Eurobarometer 11 (Apr 1979)*. Cologne: GESIS Data Archive, ZA1036 Data file Version 1.0.1.

Conover, P. J. and S. Feldman (1982). Projection and the Perception of Candidates' Issue Positions. *Political Research Quarterly 35*(2), 228–244.

Converse, P. E. (1964). The Nature of Belief Systems in Mass Publics. In D. Apter (Ed.), *Ideology and Discontent*, pp. 206–261. New York: Free Press.

Cox, D. R. (1970). *The Analysis of Binary Data*. London: Methuen.

Cox, D. R. (1981). Statistical issues in the analysis of time series: some recent developments. *Scandinavian Journal of Statistics 8*(2), 93–115.

Cox, G. W. (1984). Electoral equilibrium in double member districts. *Public Choice 44*(3), 443–451.

Cramer, J. S. (2005). Omitted Variables and Misspecified Disturbances in the Logit Model. Tinbergen Institute Discussion Paper 05-084/4, Amsterdam and Rotterdam.

Croissant, Y. (2012). *Estimation of multinomial logit models in R: The mlogit Package*. R package version 0.2-2. URL: http://cran. r-project. org/web/packages/mlogit/vignettes/mlogit.pdf.

Dahl, R. A. (1971). *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.

Davis, O. A., M. J. Hinich, and P. C. Ordeshook (1970). An Expository Development of a Mathematical Model of the Electoral Process. *American Political Science Review 64*(2), 426–448.

de Vries, C. E. (2007). Sleeping Giant: Fact or Fairytale? How European Integration Affects National Elections. *European Union Politics 8*(3), 363–385.

de Vries, C. E. (2010). EU Issue Voting: Asset or Liability? How European Integration Affects Parties' Electoral Fortunes. *European Union Politics 11*(1), 89–117.

de Vries, C. E. and S. B. Hobolt (2012). When dimensions collide: The electoral success of issue entrepreneurs. *European Union Politics 13*(2), 246–268.

de Vries, C. E. and S. B. Hobolt (2016). Public Support for European Integration. *Annual Review of Political Science 19*, 22.1–22.20.

de Vries, C. E. and G. Marks (2012). The struggle over dimensionality: A note on theory and empirics. *European Union Politics 13*(2), 185–193.

de Vries, C. E. and E. R. Tillman (2011). European Union issue voting in East and West Europe: The role of political context. *Comparative European Politics 9*(1), 1–17.

de Vries, C. E., W. van der Brug, M. H. van Egmond, and C. van der Eijk (2011). Individual and contextual variation in EU issue voting: The role of political information. *Electoral Studies 30*(1), 16–28.

Denzau, A. T. and R. J. Mackay (1981). Structure-Induced Equilibria and Perfect-Foresight Expectations. *American Journal of Political Science 25*(4), 762–779.

Diggle, P., P. Heagerty, K.-Y. Liang, and S. Zeger (1994). *Analysis of Longitudinal Data.* Oxford: Oxford University Press.

Dinkel, R. (1977). Der Zusammenhang zwischen Bundes- und Landtagswahlergebnissen. *Politische Vierteljahresschrift 18*(3), 348–360.

Dobson, D. and D. A. Meeter (1974). Alternative Markov Models for Describing Change in Party Identification. *American Journal of Political Science 18*(3), 487.

Dow, J. K. and J. W. Endersby (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 Political Action in a Democracy. *Journal of Political Economy 65*(2), 135–150.

Dunning, T. (2008). Improving Causal Inference: Strengths and Limitations of Natural Experiments. *Political Research Quarterly 61*(2), 282–293.

Dunning, T. (2012). *Natural Experiments in the Social Sciences: A Design-Based Approach.* Cambridge: Cambridge University Press.

Enelow, J. M. and M. J. Hinich (1984a). Probabilistic Voting and the Importance of Centrist Ideologies in Democratic Elections. *The Journal of Politics 46*(2), 459–478.

Enelow, J. M. and M. J. Hinich (1984b). *The Spatial Theory of Voting: An Introduction.* New York: Cambridge University Press.

Epstein, D. L., R. Bates, J. Goldstone, I. Kristensen, and S. O'Halloran (2006). Democratic Transitions. *American Journal of Political Science 50*(3), 551–569.

Erdem, T. (1996). A Dynamic Analysis of Market Structure Based on Panel Data. *Marketing Science 15*(4), 359–378.

Evans, G. (1998). Euroscepticism and Conservative Electoral Support: How an Asset Became a Liability. *British Journal of Political Science 28*(4), 573–590.

Evans, G. (2002). European Integration, Party Politics and Voting in the 2001 Election. *British Elections and Parties Review 12*(1), 95–110.

Evans, G. and R. Andersen (2006). The political conditioning of economic perceptions. *Journal of Politics 68*(1), 194–207.

Faas, T. and H. Schoen (2006). The importance of being first: Effects of candidates' list positions in the 2003 Bavarian state election. *Electoral Studies 25*(1), 91–102.

Fauvelle-Aymar, C. and A. François (2015). Mobilization, cost of voting and turnout: a natural randomized experiment with double elections. *Public Choice 162*(1), 183–199.

Fauvelle-Aymar, C. and M. Stegmaier (2008). Economic and political effects on European Parliamentary electoral turnout in post-communist Europe. *Electoral Studies 27*(4), 661–672.

Feld, S. L., S. Merrill III, and B. Grofman (2014). Modeling the effects of changing issue salience in two-party competition. *Public Choice 158*(3-4), 465–482.

Ferrara, F., E. S. Herron, and M. Nishikawa (2005). *Mixed Electoral Systems.* Basingstoke: Palgrave Macmillan.

Ferrara, F. and J. T. Weishaupt (2004). Get your Act Together. *European Union Politics 5*(3), 283–306.

Ferwerda, J. (2014). Electoral consequences of declining participation: A natural experiment in Austria. *Electoral Studies 35*, 242–252.

Finke, D. (2009). Estimating the Effect of Nonseparable Preferences in EU Treaty Negotiations. *Journal of Theoretical Politics 21*(4), 543–569.

Finke, D. and A. Fleig (2013). The merits of adding complexity: non-separable preferences in spatial models of European Union politics. *Journal of Theoretical Politics 25*(4), 546–575.

Fiorina, M. P. (1996). *Divided Government.* Boston, MA: Allyn and Bacon.

Fiorina, M. P., S. A. Abrams, and J. C. Pope (2008). Polarization in the American Public: Misconceptions and Misreadings. *The Journal of Politics 70*(2), 556–560.

Flickinger, R. S. and D. T. Studlar (2007). One Europe, Many Electorates? Models of Turnout in European Parliament Elections After 2004. *Comparative Political Studies 40*(4), 383–404.

Franklin, M. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945.* Cambridge: Cambridge University Press.

Franklin, M. N. and S. B. Hobolt (2011). The legacy of lethargy: How elections to the European Parliament depress turnout. *Electoral Studies 30*(1), 67–76.

Gabel, M. (2000). European integration, voters and national politics. *West European Politics 23*(4), 52–72.

Gabel, M. J. and C. J. Anderson (2004). The structure of citizen attitudes and the European political space. In G. Marks and M. R. Steenbergen (Eds.), *European Integration and Political Conflict*, pp. 13–32. Cambridge: Cambridge University Press.

Gallagher, M. and P. Mitchell (2005). *The Politics of Electoral Systems.* Oxford: Oxford University Press.

Garry, J. and J. Tilley (2015). Inequality, state ownership and the European Union: How economic context and economic ideology shape support for the European Union. *European Union Politics 16*(1), 139–154.

Gelman, A. (2008). *Red state, blue state, rich state, poor state: Why Americans vote the way they do.* Princeton: Princeton University Press.

Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research. *Electoral Studies 25*(4), 637–663.

Giebler, H. and A. Wagner (2015). Contrasting First- and Second-Order Electoral Behaviour: Determinants of Individual Party Choice in European and German Federal Elections. *German Politics 24*(1), 46–66.

Golder, S. N., I. Lago, A. Blais, E. Gidengil, and T. Gschwend (2017). *Multi-Level Electoral Politics. Beyond the Second-Order Election Model.* Oxford: Oxford University Press.

Granberg, D. and S. Holmberg (1986). Political Perception among Voters in Sweden and the U. S.: Analyses of Issues with Explicit Alternatives. *Western Political Quarterly 39*(1), 7–28.

Grand, P. and G. Tiemann (2013). Projection effects and specification bias in spatial models of European Parliament elections. *European Union Politics 14*(4), 497–521.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal 7*(1), 98–119.

Gschwend, T. (2007). Ticket-splitting and strategic voting under mixed electoral rules: Evidence from Germany. *European Journal of Political Research 46*(1), 1–23.

Gschwend, T., L. F. Stoetzer, and S. Zittlau (2016). What drives rental votes? How coalitions signals facilitate strategic coalition voting. *Electoral Studies 44*, 293–306.

Habermas, J. (2012). *The Crisis of the European Union: A Response*. Cambridge: Polity Press.

Halaby, C. N. (2004). Panel Models in Sociological Research: Theory into Practice. *Annual Review of Sociology 30*, 507–544.

Harvey, A. C. (1978). The Estimation of Time-Varying Parameters from Panel Data. *Annales de l'inséé 30/31*, 203–226.

Hausman, J., J. Abrevaya, and F. Scott-Morton (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics 87*(2), 239–269.

Heath, A., I. McLean, B. Taylor, and J. Curtice (1999). Between first and second order: a comparison of voting behaviour in european and local elections in britain. *European Journal of Political Research 35*(3), 389–414.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153.

Heckman, J. J. (1981). Statistical models for discrete panel data. In C. F. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 2–50. Cambridge, MA: MIT Press.

Henningsen, A. and O. Toomet (2010). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics 26*(3), 443–458.

Herron, E. S. and M. Nishikawa (2001). Contamination effects and the number of parties in mixed-superposition electoral systems. *Electoral Studies 20*(1), 63–86.

Herron, M. (1999). Postestimation Uncertainty in Limited Dependent Variables Models. *Political Analysis 8*(1), 83–98.

Hillygus, D. S. (2007). The Dynamics of Voter Decision Making Among Minor-Party Supporters: The 2000 Presidential Election in the United States. *British Journal of Political Science 37*(02), 225–244.

Hillygus, D. S. and S. Jackman (2003). Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy. *American Journal of Political Science 47*(4), 583.

Hinich, M. J. and M. C. Munger (1997). *Analytical Politics.* Cambridge, New York, Melbourne: Cambridge University Press.

Hinich, M. J. and P. C. Ordeshook (1969). Abstentions and equilibrium in the electoral process. *Public Choice 7*(1), 81–106.

Hirschman, A. O. (1970). *Exit, Voice and Loyalty. Responses to Decline in Firms, Organizations and States.* Cambridge, MA: Harvard University Press.

Hix, S. (1999). *The political system of the European Union.* Houndmills: Macmillan.

Hix, S. and C. Lord (1997). *Political parties in the European Union.* Basingstoke: Macmillan.

Hix, S. and M. Marsh (2007). Punishment or Protest? Understanding European Parliament Elections. *The Journal of Politics 69*(2), 495–510.

Hix, S. and M. Marsh (2011). Second-order effects plus pan-European political swings: An analysis of European Parliament elections across time. *Electoral Studies 30*(1), 4–15.

Hobolt, S. B. (2014). A vote for the President? The role of Spitzenkandidaten in the 2014 European Parliament elections. *Journal of European Public Policy 21*(10), 1528–1540.

Hobolt, S. B. (2015). The 2014 European Parliament Elections: Divided in Unity? *JCMS: Journal of Common Market Studies 53*(5), 6–21.

Hobolt, S. B. and C. E. de Vries (2016). Turning against the Union? The impact of the crisis on the Eurosceptic vote in the 2014 European Parliament elections. *Electoral Studies 44*, 504–514.

Hobolt, S. B. and J.-J. Spoon (2012). Motivating the European voter: Parties, issues and campaigns in European Parliament elections. *European Journal of Political Research 51*(6), 701–727.

Hobolt, S. B., J.-J. Spoon, and J. Tilley (2009). A Vote Against Europe? Explaining Defection at the 1999 and 2004 European Parliament Elections. *British Journal of Political Science 39*(1), 93–115.

Hobolt, S. B. and J. Wittrock (2011). The second-order election model revisited: An experimental test of vote choices in European Parliament elections. *Electoral Studies 30*(1), 29–40.

Hooghe, L. and G. Marks (2001). *Multi-Level Governance and European Integration.* Lanham, MD: Rowman and Littlefield.

Hooghe, L., G. Marks, and C. J. Wilson (2002). Does Left/Right Structure Party Positions on European Integration? *Comparative Political Studies 35*(8), 965–989.

Jackman, S. (1999). Correcting surveys for non-response and measurement error using auxiliary information. *Electoral Studies 18*(1), 7–27.

Jackman, S. (2000). In and Out of War and Peace: Transitional Models of International Conflict. *Unpublished Manuscript*, 1–32.

Jankowski, M., S. Schneider, and M. Tepe (2016). Ideological alternative? Analyzing Alternative für Deutschland candidates' ideal points via black box scaling. *Party Politics*, 1–13.

Jessee, S. A. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review 103*(1), 59–81.

Johnston, R., R. Sarker, K. Jones, A. Bolster, C. Propper, and S. Burgess (2005). Egocentric economic voting and changes in party choice: Great Britain 1992–2001. *Journal of Elections, Public Opinion & Parties 15*(1), 129–144.

Kadane, J. B. (1972). On division of the question. *Public Choice 13*(1), 47–54.

Katz, E. (2001). Bias in Conditional and Unconditional Fixed Effects Logit Estimation. *Political Analysis 9*(4), 379–384.

Katz, J. N. and G. Katz (2010). Correcting for survey misreports using auxiliary information with an application to estimating turnout. *American Journal of Political Science 54*(3), 815–835.

Katz, R. S. and P. Mair (1995). Changing Models of Party Organization and Party Democracy: The Emergence of the Cartel Party. *Party Politics 1*(1), 5–28.

Keane, M. P. (2013). Panel Data Discrete Choice Models of Consumer Demand. *Unpublished Manuscript*.

Kedar, O. (2005). When Moderate Voters Prefer Extreme Parties: Policy Balancingin Parliamentary Elections. *American Political Science Review 99*(2), 185–199.

Kedar, O. (2006). How voters work around institutions: Policy balancing in staggered elections. *Electoral Studies 25*(3), 509–527.

Kießling, A. (2008). Das Parteiensystem Bayerns. In U. Jun, M. Haas, and O. Niedermayer (Eds.), *Parteien und Parteiensysteme in den deutschen Ländern*. Wiesbaden: VS, Verlag für Sozialwissenschaften.

King, G., R. O. Keohane, and S. Verba (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

King, G., M. Tomz, and J. Wittenberg (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science 44*(2), 347.

Klar, S. (2014). A Multidimensional Study of Ideological Preferences and Priorities among the American Public. *Public Opinion Quarterly 78*, 344–359.

Kousser, T. (2004). Retrospective voting and strategic behavior in European Parliament elections. *Electoral Studies 23*(1), 1–21.

Krailo, M. D. and M. C. Pike (1984). Algorithm AS 196: Conditional Multivariate Logistic Analysis of Stratified Case-Control Studies. *Applied Statistics 33*(1), 95–103.

Kramer, G. H. (1972). Sophisticated voting over multidimensional choice spaces. *The Journal of Mathematical Sociology 2*(2), 165–180.

Lacy, D. (2001a). A Theory of Nonseparable Preferences in Survey Responses. *American Journal of Political Science 45*(2), 239–258.

Lacy, D. (2001b). Nonseparable Preferences, Measurement Error, and Unstable Survey Responses. *Political Analysis 9*(2), 95–115.

Lacy, D. and E. M. Niou (2000). A Problem with Referendums. *Journal of Theoretical Politics 12*(1), 5–31.

Lacy, D. and E. M. S. Niou (1998). Elections in Double-Member Districts with Nonseparable Voter Preferences. *Journal of Theoretical Politics 10*(1), 89–110.

Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.

Lee, L.-f. (1980). Specification Error in Multinomial Logit Models: Analysis of the Omitted Variable Bias. Discussion Paper No. 80 - 131. Center for Economic Research, University of Minnesota.

Lee, M.-J. (2012). Practical Panel Conditional Logit Estimators for Binary, Ordered Discrete and Multinomial Responses. *Unpublished Manuscript*.

Lee, M.-J. (2015). Panel Conditional and Multinomial Logit Estimators. In B. H. Baltagi (Ed.), *The Oxford Handbook of Panel Data*, pp. 203–232. Oxford University Press.

Lefevere, J. and P. van Aelst (2014). First-order, second-order or third-rate? A comparison of turnout in European, local and national elections in the Netherlands. *Electoral Studies 35*, 159–170.

Leininger, A., L. Rudolph, and S. Zittlau (2016). How to Increase Turnout in Low-Salience Elections: Quasi-Experimental Evidence on the Effect of Concurrent Second-Order Elections on Political Participation. *Political Science Research and Methods*, 1–18.

Levendusky, M. S. (2009). The microfoundations of mass polarization. *Political Analysis 17*(2), 162–176.

Lewis-Beck, M. S., R. Nadeau, and A. Elias (2008). Economics, party, and the vote: Causality issues and panel data. *American Journal of Political Science 52*(1), 84–95.

Lijphart, A. (1994). *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945–1990.* Oxford: Oxford University Press.

Lin, T.-m., J. M. Enelow, and H. Dorussen (1999). Equilibrium in multicandidate probabilistic spatial voting. *Public Choice 98*(1-2), 59–82.

Lipset, S. M. and S. Rokkan (1967). *Cleavage structures, party systems, and voter alignments.* New York: Free Press.

Lo, J., S.-O. Proksch, and T. Gschwend (2013). A Common Left-Right Scale for Voters and Parties in Europe. *Political Analysis 22*(2), 205–223.

Lupia, A. (2002). New Ideas in Experimental Political Science. *Political Analysis 10*(4), 319–324.

Lupia, A. and M. D. McCubbins (1998). *The Democratic Dilemma: Can Citizens Learn What They Need to.* New York: Cambridge University Press.

Mackay, R. J. and C. L. Weaver (1981). Agenda control by budget maximizers in a multi-bureau setting. *Public Choice 37*(3), 447–472.

Maier, M., J. Maier, A. Baumert, N. Jahn, S. Krause, and S. Adam (2015). Measuring citizens' implicit and explicit attitudes towards the European Union. *European Union Politics 16*(3), 369–385.

Majone, G. (2005). *Dilemmas of European integration: The ambiguities and pitfalls of integration by stealth.* Oxford: Oxford University Press.

Manksi, C. F. (1995). *Identification Problems in the Social Sciences.* Cambridge: Harvard University Press.

Marks, G. and L. Hooghe (2009). A Postfunctionalist Theory of European Integration: From Permissive Consensus to Constraining Dissensus. *British Journal of Political Science 39*(1), 1–23.

Marks, G., M. Steenbergen, and L. Hooghe (2012). To dichotomize or not to dichotomize? *European Union Politics 13*(2), 334–339.

Marsh, M. (1998). Testing the Second-Order Election Model after Four European Elections. *British Journal of Political Science* (4), 591–607.

Marsh, M. (2003). Theories of less important elections: explanations of electoral behaviour in European parliament elections. *Revue de la Maison Francaise d'Oxford 1*(1), 189–210.

Marsh, M. (2007). European Parliament elections and losses by governing parties. In W. van der Brug, C. van der Eijk, and M. Franklin (Eds.), *European Elections & Domestic Politics: Lessons from the Past and Scenarios for the Future*, pp. 51–73. Notre Dame: University of Notre Dame Press.

Marsh, M. and S. Mikhaylov (2010). European Parliament elections and EU governance. *Living Reviews in European Governance 5*(4), 1–30.

Massicotte, L. (2011). Mixed Systems. In J. M. Colomer (Ed.), *Personal Representation. The Neglected Dimension of Electoral Systems.*, pp. 97–115. Colchester: ECPR.

Mattila, M. (2003). Why bother? Determinants of turnout in the European elections. *Electoral studies 22*(3), 449–468.

McCarty, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches.* Cambridge, MA: The MIT Press.

McDermott, R. (2002). Experimental Methods in Political Science. *Annual Review of Political Science 5*(1), 31–61.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McKelvey, R. D. (1986). Covering, Dominance, and Institution-Free Properties of Social Choice. *American Journal of Political Science 30*(2), 283–314.

Merrill III, S. and J. Adams (2001). Computing Nash Equilibria in Probabilistic , Multiparty Spatial Models with Nonpolicy Components. *Political Analysis 9*(4), 347–361.

Merrill III, S., B. Grofman, and J. Adams (2001). Assimilation and contrast effects in voter projections of party locations: Evidence from Norway, France, and the USA. *European Journal of Political Research 40*(2), 199–221.

Miller, W. L. and M. Mackie (1973). The Electoral Cycle and the Asymmerty of Government and Opposition Popularity: An Alternative Model of the Relationship between Economic Conditions and Political Popularity. *Political Studies 21*(3), 263–279.

Milyo, J. (2000). Logical deficiencies in spatial models: a constructive critique. *Public Choice 105*(3–4), 273–289.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review 26*(1), 67–82.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica 46*(1), 69–85.

Neuhaus, J. (1999). Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika 86*(4), 843–855.

Neyman, J. and E. L. Scott (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica 16*(1), 1–32.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica 49*(6), 1417–1426.

Nordhaus, W. D. (1975). The Political Business Cycle. *The Review of Economic Studies 42*(2), 169–190.

Oppenhuis, E., C. van der Eijk, and M. N. Franklin (1996). The party context: Outcomes. In C. van der Eijk and M. N. Franklin (Eds.), *Choosing Europe? The European Electorate and National Politics in the Face of Union*, pp. 287–306. Ann Arbor: University of Michigan Press.

Ordeshook, P. C. (1986). *Game theory and political theory.* New York: Cambridge University Press.

Orford, S., C. Rallings, M. Thrasher, and G. Borisyuk (2009). Electoral salience and the costs of voting at national, sub-national and supra-national elections in the UK: a case study of Brent, UK. *Transactions of the Institute of British Geographers 34*(2), 195–214.

Pawitan, Y. (2013). *In All Likelihood: Statistical Modelling and Inference Using Likelihood.* Oxford: Oxford University Press.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys 3*, 96–146.

Persson, M. (2014). Does Survey Participation Increase Voter Turnout? Reexamining the Hawthorne Effect in the Swedish National Election Studies. *Political Science Research and Methods 2*(2), 297–307.

Persson, M., M. Solevid, and R. Öhrvall (2013). Voter turnout and political equality: Testing the 'law of dispersion' in a Swedish natural experiment. *Politics 33*(3), 172–184.

Pforr, K. (2013). *femlogit : Implementation und Anwendung der multinominalen logistischen Regression mit "fixed effects "*. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

Pierce, R. (1999). Mass-Elite Issue Linkages and the Responsible Party Model of Representation. In W. E. Miller, R. Pierce, J. Thomassen, R. Herrera, S. Holmberg, P. Esaisson, and B. Webels (Eds.), *Policy Representation in Western Democracies*, pp. 9–32. Oxford: Oxford University Press.

Pinheiro, J. C. and D. M. Bates (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing 6*(3), 289–296.

Poole, K. T. and H. Rosenthal (1985). A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science 29*(2), 357–384.

Poole, K. T. and H. Rosenthal (2007). *Ideology & Congress*. New Brunswick, NJ: Transaction Publishers.

Proksch, S.-O. and J. Lo (2012). Reflections on the European integration dimension. *European Union Politics 13*(2), 317–333.

Prosser, C. (2016). Second Order Electoral Rules and National Party Systems: The Duvergerian Effects of European Parliament Elections. *European Union Politics 17*(3), 366–386.

Quinn, K. M., A. D. Martin, and A. B. Whitford (1999). Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models. *American Journal of Political Science 43*(4), 1231–1247.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.

Reif, K. (1997). European elections as member state second-order elections revisited. *European Journal of Political Research 31*, 115–124.

Reif, K. and H. Schmitt (1980). Nine Second-Order National Elections - a Conceptual Framework for the Analysis of European Election Results. *European Journal of Political Research 8*(1), 3–44.

Riker, W. H. (1982). *Liberalism against populism: A confrontation between the theory of democracy and the theory of social choice*. San Francisco: Freeman.

Rivers, D. (1988). Heterogeneity in Models of Electoral Choice. *American Journal of Political Science 32*(3), 737–757.

Rohrschneider, R. and S. Whitefield (2016). Responding to Growing European Union-Skepticism? The Stances of Political Parties toward European Integration in Western and Eastern Europe following the Financial Crisis. *European Union Politics 17*(1), 138–161.

Rubin, D. B. (1978). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics 6*(1), 34–58.

Rudolph, L. and T. Däubler (2016). Holding Individual Representatives Accountable: The Role of Electoral Systems. *Journal of Politics 78*(3), 746–762.

Sanz, A. (2008). Split-Ticket Voting in Multi-Level Electoral Competition: European, National and Regional Concurrent Elections in Spain. In C. van der Eijk and H. Schmitt (Eds.), *Multi-level electoral systems of the European Union*, pp. 101–136. Mannheim: CONNEX Report Series 04.

Schäfer, C. and M. Debus (2015). *Euroscepticism, Policy Alienation and Abstention in the 2014 EP Elections*. Paper prepared for the conference "The 2014 European Parliament Elections in Perspective: European Elections Studies 1979-2014" at the MZES, University of Mannheim (November 6-8, 2015). Unpublished Manuscript.

Schakel, A. H. (2015). How to analyze second-order election effects? A refined second-order election model. *Comparative European Politics 13*(6), 636–655.

Schakel, A. H. and C. Jeffery (2013). Are Regional Elections really 'Second-Order' Elections? *Regional Studies 47*(3), 323–341.

Schattschneider, E. E. (1960). *The semisovereign people: A realist's view of democracy in America.* New York: Holt, Rinehart and Winston.

Schmitt, H. (2005). The European Parliament Elections of June 2004: Still Second-Order? *West European Politics 28*(3), 650–679.

Schmitt, H., S. B. Hobolt, S. A. Popa, E. Teperoglou, E. Parliament, and P. M. U. Directorate-General for Communication (2016). European Parliament Election Study 2014, Voter Study, First Post-Election Survey.

Schmitt, H. and R. Mannheimer (1991). About voting and non-voting in the European elections of June 1989. *European Journal of Political Research 19*(1), 31–54.

Schmitt, H., A. Sanz, and D. Braun (2008). The Micro-Foundations of Second-order Election Theory: Theoretical Reconstructions and Empirical Tests Based Upon the European Election Study 2004. Working paper.

Schmitt, H. and C. van der Eijk (2007). Non-voting in European Parliament elections and support for European integration. In W. van der Brug and C. van der Eijk (Eds.), *European Elections & Domestic Politics: Lessons from the Past and Scenarios for the Future*, pp. 145–167. Notre Dame: University of Notre Dame Press.

Schofield, N. (1978). Instability of simple dynamic games. *Review of Economic Studies 45*(3), 575–594.

Schofield, N., A. D. Martin, K. M. Quinn, and A. B. Whitford (1998). Multiparty electoral competition in the Netherlands and Germany: A model based on multinomial probit. *Public Choice 97*(2), 257–293.

Schofield, N. and I. Sened (2005). Multiparty Competition in Israel, 1988–96. *British Journal of Political Science 35*(4), 635.

Schofield, N. and A. Zakharov (2009). A stochastic model of the 2007 Russian Duma election. *Public Choice 142*(1-2), 177–194.

Schwartz, T. (1977). Collective Choice, Separation of Issues and Vote Trading. *The American Political Science Review 71*(3), 999–1010.

Sekhon, J. S. and R. Titiunik (2012). When Natural Experiments Are Neither Natural nor Experiments. *American Political Science Review 106*(1), 35–57.

Selb, P. and S. Munzert (2013). Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys. *Electoral Studies 32*(1), 186–196.

Shachar, R. and Z. Eckstein (2007). Correcting for bias in retrospective data. *Journal of Applied Econometrics 22*(3), 657–675.

Shafer, B. E. and W. J. M. Claggett (1995). *The Two Majorities: The Issue Context of American Politics*. Baltimore: Johns Hopkins University Press.

Shikano, S. (2008). Die Eigendynamik zur Eindimensionalität des Parteienwettbewerbs: Eine Simulationsstudie. *Politische Vierteljahresschrift 49*(2), 229–250.

Shugart, M. S. (2001). Electoral "efficiency" and the move to mixed-member systems. *Electoral Studies 20*(2), 173–193.

Signorino, C. S. and K. Yilmaz (2003). Strategic misspecification in regression models. *American Journal of Political Science 47*(3), 551–566.

Silver, B. D., B. A. Anderson, and P. R. Abramson (1986). Who Overreports Voting? *American Political Science Review 80*(2), 613–624.

Sitter, N. (2001). The politics of opposition and European integration in Scandinavia: Is Euroscepticism a government opposition dynamic? *West European Politics 24*(4), 22–39.

Spoon, J.-J. (2012). How salient is Europe? An analysis of European election manifestos, 1979-2004. *European Union Politics 13*(4), 558–579.

Steenbergen, M. R. and G. Marks (2004). Introduction: Models of political conflict in the European Union. In G. Marks and M. R. Steenbergen (Eds.), *European Integration and Political Conflict*, pp. 1–12. Cambridge: Cambridge University Press.

Stockemer, D. (2012). Citizens' support for the European Union and participation in European Parliament elections. *European Union Politics 13*(1), 26–46.

Stoetzer, L. F. and S. Zittlau (2015). Multidimensional Spatial Voting with Non-separable Preferences. *Political Analysis 23*(3), 415–428.

Stoetzer, L. F., S. Zittlau, T. Gschwend, and T. Witt (2015). Leihstimmen im Bundestagswahljahr 2013. *Politische Psychologie* (1), 88–107.

Studlar, D. T., R. Flickinger, and S. Bennett (2003). Turnout in European parliament elections: Towards a European-centred model. *Journal of Elections, Public Opinion & Parties 13*(1), 195–209.

Swaddle, K. and A. Heath (1989). Official and Reported Turnout in the British General Election of 1987. *British Journal of Political Science 19*(4), 537–570.

Taagepera, R. (2002). Nationwide threshold of representation. *Electoral Studies 21*(3), 383–401.

Tillman, E. R. (2004). The European Union at the Ballot Box? European Integration and Voting Behavior in the New Member States. *Comparative Political Studies 37*(5), 590–610.

Treier, S. and D. S. Hillygus (2009). The nature of political ideology in the contemporary electorate. *Public Opinion Quarterly 73*(4), 679–703.

Tsebelis, G. and G. Garrett (2000). Legislative Politics in the European Union. *European Union Politics 1*(1), 9–36.

Tufte, E. R. (1975). Determinants of the Outcomes of Midterm Congressional Elections. *American Political Science Review 69*(3), 812–826.

van Aelst, P. and J. Lefevere (2011). Has Europe got anything to do with the European elections? A study on split-ticket voting in the Belgian regional and European elections of 2009. *European Union Politics 13*(1), 3–25.

van de Wardt, M., C. E. de Vries, and S. B. Hobolt (2014). Exploiting the Cracks: Wedge Issues in Multiparty Competition. *Journal of Politics 76*(4), 986–999.

van der Eijk, C. and M. N. Franklin (1991). European Community politics and electoral representation: Evidence from the 1989 European Elections Study. *European Journal of Political Research 19*(1), 105–127.

van der Eijk, C. and M. N. Franklin (1996). *Choosing Europe? The European Electorate and National Politics in the Face of Union.* Ann Arbor: University of Michigan Press.

van der Eijk, C. and M. N. Franklin (2004). Potential for contestation on European Matters at national election in Europe. In G. Marks and M. R. Steenbergen (Eds.), *European Integration and Political Conflict*, pp. 32–51. Cambridge, MA: Cambridge University Press.

van der Eijk, C. and B. Niemöller (1979). Recall accuracy and its determinants. *Acta Politica 14*(3), 289–342.

van der Eijk, C. and H. Schmitt (2008). Multi-level electoral systems of the European Union: elaborating existing approaches and defining the research agenda for the future. In C. van der Eijk and H. Schmitt (Eds.), *The Multilevel Electoral System of the EU*, pp. 1–20. Mannheim: CONNEX Report Series 04.

van Egmond, M., W. van der Brug, S. Hobolt, M. Franklin, and E. V. Sapir (2013). *European Parliament Election Study 2009, Voter Study.* Cologne: GESIS Data Archive, ZA5055 Data file Version 1.1.0.

van Spanje, J. and C. de Vreese (2011). So what's wrong with the EU? Motivations underlying the Eurosceptic vote in the 2009 European elections. *European Union Politics 12*(3), 405–429.

Waldahl, R. and B. O. Aardal (2000). The Accuracy of Recalled Previous Voting: Evidence from Norwegian Election Study Panels. *Scandinavian Political Studies 23*(4), 373–389.

Weber, T. (2007). Campaign Effects and Second-Order Cycles: A Top-Down Approach to European Parliament Elections. *European Union Politics 8*(4), 509–536.

Weber, T. (2011). Exit, Voice, and Cyclicality: A Micrologic of Midterm Effects in European Parliament Elections. *American Journal of Political Science 55*(4), 907–922.

Weir, B. T. (1975). The Distortion of Voter Recall. *American Journal of Political Science 19*(1), 53–62.

Whitten, G. D. and H. D. Palmer (1996). Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands. *American Journal of Political Science 40*(1), 231–260.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: The MIT press.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data: Second Edition.* Cambridge, MA: The MIT Press.

Xie, Y. (2013). Population heterogeneity and causal inference. *Proceedings of the National Academy of Sciences 110*(16), 6262–6268.

Yatchew, A. and Z. Griliches (1985). Specification Error in Probit Models. *The Review of Economics and Statistics 67*(1), 134–139.