MANNHEIMER
ZENTRUM FÜR
EUROPÄISCHE
SOZIALFORSCHUNG

Policy-Specific Alienation and Indifference in the Calculus of Voting:

A Simultaneous Model of Party Choice and Abstention

Paul W. Thurner, Angelika Eymann

Arbeitsbereich II / Nr. 18
Mannheim 1997
ISSN 0948-0080

Working papers

Arbeitspapiere

Paul W. Thurner and Angelika Eymann

Policy-specific Alienation and Indifference in the Calculus of Voting:

A Simultaneous Model of Party Choice and Abstention

Paul W. Thurner, Angelika Eymann:
Policy-specific Alienation and Indifference in the Calculus of Voting : A Simultaneous Model of Party Choice and Abstention.
Mannheimer Zentrum für Europäische Sozialforschung (MZES).
Mannheim, 1997 (revised version).
(Arbeitspapiere Arbeitsbereich II / 18)
ISSN 0948-0080

Deckblattlayout: Uwe Freund

Nicht im Buchhandel erhältlich
Schutzgebühr: DM 5,--
Bezug:
Mannheimer Zentrum für Europäische Sozialforschung (MZES), Postfach, 68131 Mannheim

Editorial Note:

Paul W. Thurner is research fellow at the Mannheim Centre for European Social Research (MZES). His research topics are the modelling of voting in multiparty systems as well as international negotiations and institutional choice.
(E-mail: paul.thurner@mzes.uni-mannheim.de)
Angelika Eymann is "wissenschaftliche Assistentin" at the department of economics at the University of Mannheim. Describing herself as an applied microeconometrician, she currently concentrates on the empirical analysis of the portfolio choice of German households. Previous studies focused on tourism demand and the EU's trade policy.
(E-mail: angelika@econ.uni-mannheim.de)

Redaktionelle Notiz:

Paul W. Thurner ist wissenschaftlicher Mitarbeiter am Mannheimer Zentrum für Europäische Sozialforschung (MZES). Seine Forschungsschwerpunkte sind die Modellierung von Wählerverhalten im Mehrparteiensystem sowie kollektive Entscheidungen mit dem Schwerpunkt Institutionenbildung in internationalen Verhandlungen.
Angelika Eymann ist wissenschaftliche Assistentin am Lehrstuhl für Makroökonomik und Wirtschaftspolitik der Universität Mannheim. Die Analyse des Anlageverhaltens deutscher Haushalte bildet zur Zeit den Schwerpunkt ihrer Forschungsarbeiten im Bereich der angewandten Mikroökonometrie. Frühere Arbeiten bezogen sich auf die Tourismus-nachfrage sowie die Außenhandelspolitik der EU.

Contents:

Abstract

This article aims to combine the spatial models of candidate/party choice and abstention/participation choice in a single decision model and to provide empirical reaction functions. The starting point of this study is the hypothesis of a sequential ordering of candidate choice and participation choice (Riker/Ordeshook, 1973), which has never been discussed in detail in the literature on economic theory of voting. First, the general features of hierarchical multi-stage decision making are discussed. Then an operational model for party choice and abstention choice is presented. The proposed model is more complete than previous theoretical and empirical studies since 1) it simultaneously considers the choice of n>2 parties and the abstention option; 2) it differentiates between policy-specific effects of alienation and indifference for the first time. The models are estimated with nested multiattributive discrete choice models which are theoretically founded on random utility theory.

* We are grateful to Axel Börsch-Supan, Jan van Deth, Gebhard Kirchgässner and Franz Urban Pappi for helpful comments.

Introduction

In spatial election theory, policy options offered by candidates/parties are represented by points in a finite-dimensional vector space. Each voter has a utility function on this space, which is commonly assumed to be a decreasing function of the Euclidean distance from the voter's ideal point. The spatial conceptualization of politics has given rise to a significant body of theoretical literature[1] with important, sometimes counterintuitive insights into real political processes. The paradox of not-voting has been discovered in the context of the spatial theory of voting (cf. Downs, 1957), but spatial aspects of abstention/participation have surprisingly been neglected in the enormous amount of theoretical and empirical literature on this subject. Therefore we propose to reaccentuate spatial aspects in the explanation of electoral abstention/participation, to combine the spatial models of candidate/party choice and abstention/participation choice in a single decision model and to provide empirical reaction functions.

The article proceeds as follows: First, we will shortly review the initial theoretical discussion as provided by Downs (1957) and Riker/Ordeshook (1968, 1973). Second, a short review of empirical applications of the spatial model of abstention will be given in order to survey open questions and problems. Then, we will have a closer look at the assumed sequential nature of voting and participation choice, since the theoretical implications of the suggestion by Riker/Ordeshook have never been discussed in detail. In this section an operational decision theoretic tool for the modeling of static (quasi-) hierarchical as well as multidimensional choices will be identified: McFadden's (1978, 1981) nested multinomial logit model (NMNL) or tree extreme value (TEV) model. Fourth, an operational model of the simultaneous decision to participate and to choose a party in a multiparty system will be presented. After a brief introduction of data we present empirical estimations for the simultaneous abstention choice/party choice (SACPC).

The Paradox of Not-Voting and Rational Choice

Downs (1957), one of the most important pioneers of the spatial theory of voting intensely discussed the paradox of not-voting: The conceptualization of voting as an instrumentally rational act, by which voters elect a candidate whose policies are closer to those they prefer most and which is driven only by the motivation to affect the outcome of an election, is confronted with the derived prediction that rational voters will choose to abstain. Even minimal (opportunity) costs would transcend the return of participation, since in large populations the value of one's vote is infinitesimal, the value of a vote being defined as the expected voting return as compared to the number of voters. Empirically, however, we observe large numbers of the electorate going to the ballot box. Downs therefore argued that the act of voting implies an additional long-term participation value consisting in contributing to live in a democracy (Downs, 1957: 267 ff).

Despite including a consumptive value in the calculus of voting, Downs also discussed policy-related factors leading to a propensity of abstention. He strongly insisted that abstention has to be considered as the result of the (perceived) non-discrimination of candidate positions. Contrary to Smithies' (1941) alluded analogy where abstention is described as the outcome of a voter's demand elasticity, Downs insisted that "...as long as there is even the most infinitesimal difference between (candidates/parties E./Th.) A and B, extremist voters would be forced to vote for the one closest to them, no matter how distasteful its policies appear in comparison with those of their ideal government. It is always rational per definitionem to select a greater good before a lesser, or a lesser evil before a greater" (Downs, 1957: 118 f). Indifferent citizens, that is citizens with a party differential of zero, tend to abstain.

Riker/Ordeshook (1968, 1973) systematized this theory and built the mathematical foundations[2]. They conceptualize the voting decision as a two-stage process: First, the voter decides which of the candidates to elect. In a second step he decides, whether to vote for the most preferred candidate or whether to abstain: "...we conceptualize each citizen's choices and actions as

the result of a two-stage sequential decision process. We assume, first, that the citizen evaluates both candidates and identifies a preferred candidate, and second, that he decides whether to vote or to abstain. We order the decision process in this fashion because the choice between voting and abstaining depends upon the citizen's comparative evaluation of the candidates..." (Riker/Ordeshook, 1973: 308).

The calculus of abstaining/participating has the following form (Riker/Ordeshook, 1973: 323):

$$R = PB + D - C \qquad (1)$$

A voter participates if his total expected utility R is greater than the one yielded by alternative activities, that is, if R > 0. The term B represents the related stream of utility guaranteed by the success of the more preferred party over the less preferred one. C concludes the opportunity costs arising through the voting act. P represents the probability that an individual's vote is decisive. In large electorates, this probability is very small so that even in the case of only minimal costs the term PB-C should always be negative. Riker/Ordeshook therefore introduce the D term. D measures the positive contributions to the individual's utility where the "magnitude of the effect is independent of the individual's contribution to the outcome" (Riker/Ordeshook, 1968: 27). Stigler (1972) called this utility the consumptive benefit of voting, Fiorina (1981) and Brennan/Lomasky (1993) termed it 'expressive benefits' compared to 'instrumental benefits' resulting from the party differential. Riker/Ordeshook (1968: 28) considered the following points as possible arguments to the expressive component:

- the satisfaction from compliance with the ethic of voting,
- the satisfaction from affirming allegiance to the political system,
- the satisfaction from affirming a partisan preference,

and pointed to the affirmation of a partisan preference as the most important criterion: "For many voters this must be the most important and politically significant feature of voting (Riker/Ordeshook, 1968: 28).

Despite the obvious dominance of these consumptive aspects the authors claimed that it should nevertheless be possible to formulate propositions about the interdependence of party strategies and the participation choice. They assumed that two policy-related causes are central: indifference and alienation. The harder it is to discriminate between candidates, the greater should be the probability of abstention by indifference. Alienation, on the other hand, is induced by the perceived distance to the most preferred candidate. The further the candidate moves away from the voter's ideal point, the greater the probability that he will abstain. Contrary to indifference, the alienation hypothesis is not predicted by theories of rational choice, when considering only the effect on the immediate outcome. Under the broader concept of rationality including expressive costs and benefits it can be conceived as an element of the D term, however.

It is not intended here to review the whole array of proposed theoretical solutions to the voting paradox[3]. It suffices to say that the use of the economic model of voting lies in "accounting for changes on choices, rather than choices per se" (Grofman, 1995: 94). Kirchgässner (1992) has shown convincingly in his article, that the act of voting is fully compatible with rational choice theory when considering the low cost situation of this decision: "Therefore the main difference between usual economic decisions and the voting decision is not that moral rules are not existent for economic behavior or that self-interest is not relevant for voting decisions, but that the costs of following social (moral) rules, if this implies a deviation from self-interested behavior, are usually high in economic decisions and rather low in voting decisions" (Kirchgässner, 1992: 310). It follows that the assumptions of rational choice apply also to voting: "The basic 'credo' of economic analysis, namely that relative prices matter, can hold in such situations as well. As long as our results are only concerned with behavioral changes resulting from changed restrictions, traditional results carry over at least partially into the area of low-cost-decisions". (Kirchgässner, 1992: 307). So the answer to the question: "Is turnout the paradox that ate rational

choice theory ?" (Grofman, 1995: 93) is clearly negative. What is important is to identify the ‘relative prices’ in this situation and to empirically provide reaction functions for changing conditions. The next section reviews previous empirical applications of the spatial theory of abstention.

Previous Empirical Applications of the Spatial Theory of Abstention

Despite its theoretical and formal sophistication, there are only few empirical applications of the spatial theory of voting. In a recent appreciation of Down’s ‘An Economic Theory of Democracy’, Grofman stated "...that applications of the rational choice approach to the analysis of political behavior is in its infancy" (Grofman, 1995: 13). This complaint has to be intensified even when considering the number of studies testing effects of alienation and indifference: As far as we know there are only three studies which explicitly use alienation and indifference concepts.

The first attempt to empirically test the theoretically identified arguments of the calculus of voting, including an explicit indifference term B, has been carried out by Riker/Ordeshook (1968), who used three pre-election surveys for the Presidential elections in 1952, 1956, 1960. The measure of indifference depends on the respondent’s own report how much he ‘cares’ about the outcome of an election. The authors’ argument for this operationalization is based on an additional assumption that "the greater the differential benefit from the election of his favored candidate, the more the voter can be expected to ‘care’ about the outcome". Empirical results are presented in a four-dimensional crosstabulation of dichotomized arguments: P(closeness) by B (reward) by D (civic duty) by abstention (yes/no). Indifference (B) seems to make a difference in most cells, but a straightforward interpretation of these results in terms of relative influences is rather awkward. Regarding the result of indifference the authors’ conclusion is as follows: "It is likely that B is much higher for many people than anyone has heretofore supposed" (Riker/Ordeshook, 1968: 39)

In a further study Brody/Page (1973), focusing exclusively on the effects of alienation and indifference, used two nationwide opinion surveys after the November 1968 US presidential election. In one survey, respondents were asked to indicate their evaluations of each presidential candidate on a scale

ranging from 1 to 100, in the other survey the thermometer question has been applied to evaluate the candidates on 0 to 100 ranging scale. Both were intended to elicit a summary judgment of the candidates. Bivariate tables (1973: 5-6) of alienation and turnout as well as indifference and turnout justify that both factors affect turnout.

The only article which tried to determine relative effects of alienation and indifference by means of estimation is Guttmann/Hilger/Shachmurove (1994). They raised the question whether the participation choice could be understood as an investive rather than a consumptive act. Following the argumentation of the authors, the voting act should be seen as an investment only if indifference effects are predominant. Alienation effects, on the contrary were considered as an indicator of the expressive character of participation. Using thermometer scores and applying logistic regression analysis they show empirically that participation should be conceived more as a consumptive act: The effect of alienation is smaller and insignificant (Guttmann/Hilger/Shachmurove, 1994: 203). This is considered as a piece of evidence that the importance of the voting paradox has been overestimated.

Several open questions and problems remain when considering previous empirical studies. First, until now, none of the empirical analyses has tried to utilize the ideological left-right scale or even concrete policy scales in order to determine the relative effects of policy-specific alienation and indifference upon abstention. Second, using a general single dimension, as provided by the thermometer question, does not allow to consider the possibility, that in some political debates of a campaign, alienation could be predominant, whereas in others it could be indifference that prevails. How to arrive at results in multidimensional contests? Furthermore, Riker/Ordeshook have indicated further expressive arguments as e.g. party loyalty, which has not been controlled for in the empirical analyses. The model has never been applied to multiparty system settings. Last but not least, due to the complete absence of theoretical consideration of the hypothesized ‘sequentiality’ of candidate/party choice and abstention choice, the explanation of participation has always been modeled as an independent decision. These open questions will be treated in the following.

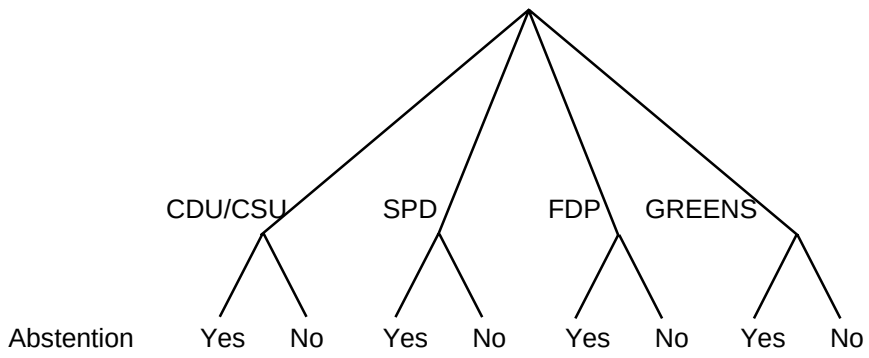
Sequential Decisions, Preference Trees and Nested Logit Model

As has been shown, Riker/Ordeshook (1973) define the voting decision as a multistage decision problem. The implications of this characterization as a sequential decision process has never been discussed in detail, however.

In general, decision problems are represented in a static manner: It is pretended that there is only one point of choice where the decision maker has to decide between actions $a_1,a_2,...a_j,..,a_J$. With the choice of one alternative, the decision is made and the decision problem resolved. Multi-stage decision problems, on the other hand, pretend that one decision problem is followed by others. Things become even more complicated, when problems are interrelated and the possible actions in later problems depend upon which actions were chosen in earlier problems. Sequentiality strictu sensu means that time passes between decisions. Such multi-stage problems can be represented as decision trees[4], with a decision point or node representing the decision maker's choice between possible actions. Combinations of subsequent, adjacent edges are called a path. The hierarchical character "....arise(s) naturally because the implementation of any decision requires a sequence of actions which are necessarily separated in time." (Pudney, 1989: 137).

If we consider only the four major parties[5] in the setting of the German multiparty system, the tree representation of the sequentiality hypothesis put forward by Riker/Ordeshook (1973) would lead to the following tree structure:

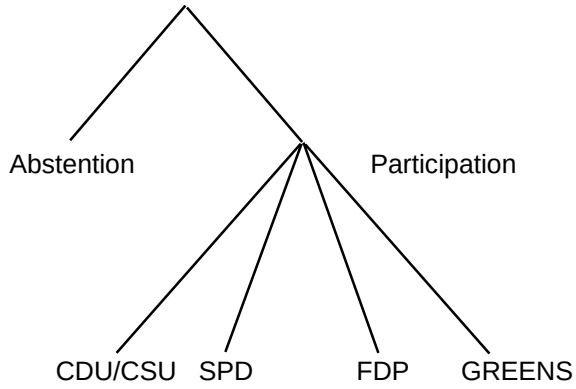
Figure 1: Party Choice and Abstention Choice Conceptualized as Decision Tree I



In a first step the voter evaluates the set of four candidates/parties and determines the best alternative. In the second stage he decides whether to abstain or to participate. Decision criteria in the first step are policy distances in the multidimensional policy space (Davis/Hinich/Ordeshook, 1970, Riker/Ordeshook, 1973), decision criteria in the second stage are alienation, indifference, and arguments of the D term, among which party loyalty has been argued to be the most important.

It is questionable, however, whether this structure of sequence is the only appropriate one. Riker/Ordeshook's argument for this type of ordering of the decision process: "Because the choice between voting and abstaining depends upon the citizen's comparative evaluation of the candidates..." (Riker/Ordeshook, 1973: 308) is not compelling since alienation and indifference imply a comparative evaluation of the candidate's positions as well. Therefore, one might also imagine the following ordering of sequences:

Figure 2: Party Choice and Abstention Choice Conceptualized as Decision Tree II

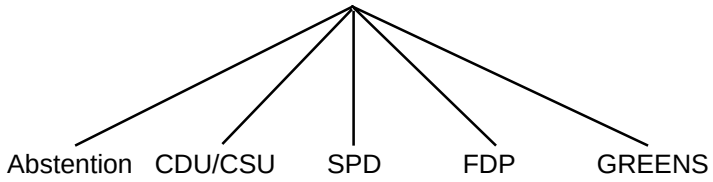


In this tree, the sequential model begins with a simple binary choice between the branches participation versus abstention. Only the second stage, if participation is chosen, involves the selection of the most preferred candidate/party.

Last but not least, it seems doubtful whether the two decisions are made at different points in time. The following representation implies that all elemental

alternatives, four candidates and the abstention option, are evaluated at the same point in time.

Figure 3: Party Choice and Abstention Choice Conceptualized as Decision Tree III



In order to clarify which of these graphical, rather intuitive representations apply to the act of voting, it is necessary to consider their respective theoretical implications.

First, it has to be noted, that additional to dynamic models, for which the hierarchical character of decisions is quasi-natural, there also exists a tradition in the literature on stochastic decision models which describes the concept of preference trees as a purely analytical tool[6]. In this context, the tree representation does not necessarily imply that an individual follows a path down the tree, making intertemporal choices in a certain order. Two approaches may be distinguished:

a) Elimination models, where the hierarchical order serves as an hypothesis on 'covert' cognitive processes (Tversky, 1972a, 1972b, Tversky/ Sattah, 1979);
b) nested logit models where the multi-level hierarchy of groupings of alternatives represents simultaneous recursive choices between subsets of alternatives with different patterns of substitution (McFadden 1978, 1981).

In the models proposed by Tversky and Tversky/Sattah, the choice of an alternative can be considered as a process in which alternatives are successively eliminated until a single one remains. Each alternative can be described by the list of characteristics it embodies.

The sequence of selecting an alternative is as follows: First, a characteristic is selected and all alternatives that do not possess this characteristic are eliminated from the choice set. A second characteristic is then selected as the

criterion for eliminating alternatives from the remaining choice set of alternatives and so on. The hierarchical elimination model (HEM) is a special type of a probabilistic choice model for which "the probability of choosing an alternative x from an offered set A is the product of the probabilities of selecting the branches containing x at each stage of the process, and the probability of selecting a branch is proportional to its overall weight" (Tversky/Sattah, 1979: 547). As McFadden (1981) has noted, in contrast to the Luce model (Luce, 1959) the elimination-by-aspects model is well suited to account for different similarities of alternatives.

However, the form of this model has also several drawbacks, preventing wider applications in economics. First, the model is restricted to characteristics that are qualitative. Second, consistent aggregation of individual choices is problematic. Third, as McFadden has pointed out, "when data are observed only on final choices, as is usual in economic applications, it is impossible to determine whether this elimination heuristic describes decision processes used by individuals" (McFadden, 1981: 231). Evidence from verbal protocols in laboratory experiments is needed in order to describe cognitive choice processes as a hierarchical elimination of aspects. One of the most important drawbacks, however, is the very restrictive view of the choice process, which "rules out any possibility of a trade off between characteristics." (Pudney, 1989: 122).

More compatible with microeconomic theory are discrete choice models. Discrete choice models, such as the nested multinomial logit model (NMNL) developed mainly by McFadden (McFadden, 1974, 1978, 1981)[7], have been explicitly derived from random utility theory. Nested logit models presume a multi-level hierarchy of groupings of alternatives that are more similar to each other internally than relative to alternatives in other subsets. Again, the ordering of such nests of alternatives can be visualized as a ‘decision tree’. A clear-cut definition of sequentiality is necessary, however. Hensher/Johnson (Hensher/Johnson, 1981) differentiate between two conditional decision structures: the sequential structure and the recursive (sequential) structure.

Following the definition of Hensher/Johnson, decisions are sequential when actors make subsequent choices conditional on the previous choices, yet

without feedback loops or mutual interactions between decisions. This form of the hierarchical structuring of the decision process implies that the utility function is additively separable:

$$U_{abc} = U_a + U_{b|a} + U_{c|ab} \text{ ,} \qquad (2)$$

with $a \in A$, $b \in B$, $c \in C$, and A, B, C being disjoint sets. For the assumed sequence a joint probability can be obtained:

$$P_{abc \in ABC} = P_{a \in A} \cdot P_{b \in B} \cdot P_{c \in C} \text{ .} \qquad (3)$$

Hensher/Johnson emphasize "that all the attributes in choice set A are independent of the attributes in B and C, the attributes in choice set B are conditioned on the selection of a particular alternative in choice set A, and the attributes in choice set C are conditioned on the selection of a particular alternative in A and in B. There is no feedback loop" (Hensher/Johnson, 1981: 80). Therefore, joint probabilities of a specific sequence of choice are simply the product of each marginal. Since there are no mutually interactive effects, joint probabilities can be determined in any order, despite the theoretical assumption of a particular sequence.

The recursive type is defined as "...an extension of the sequential structure to incorporate feedback effects, and is referred to as recursive (sequential) structure. At each decision level the choice is assumed to be conditioned on the previous choice in totality rather than a single alternative in a previous decision" (Hensher/Johnson, 1981: 81). Recursiveness in such decision trees means that different decision levels are connected with each other in a way that the attributes of all lower branch alternatives, incorporated in an additional variable, often referred to as inclusive value, expected maximum utility, or accessibility index, influence the decision among any set of upper branches. Choices on higher levels "are made in the light of the fact that the lower-level alternatives are already chosen as optimal in the respective lower subsets. This

simultaneity is the fundamental difference of the nested logit model from a pure sequence of otherwise unrelated choice models" (Börsch-Supan, 1987: 46).

The Nested Multinomial Logit Model

The nested logit model is consistent with stochastic utility maximization and is a member of the class of random utility models. Therefore, it seems to be especially well suited for modeling rational voter choices[8]. In this section we provide a brief review of the assumptions and properties of the nested logit model. For greater detail see McFadden (1978, 1981) and Ben-Akiva/Lerman (1985).

Random utility models (RUM) capture the analysts' uncertainty about voters' choice behavior. The derivation of discrete choice models from utility theory is based on a precise distinction between the behavior of the decision maker and the analysis of the researcher. Following Manski (1977), the assumption that utility is a random function "... reflects a lack of information regarding the characteristics of alternatives and/or decision-makers on the part of the observer" (Manski, 1977: 229). Instead of attempting to precisely predict an individual's choice we predict the probability with which an individual will choose an alternative. To give choice probabilities a theoretical foundation, utilities are conceptualized as random, consisting of a deterministic utility component V_{ij}, also called the systematic or representative component, and of a random component ε:

$$U_{ij} = U_{ij}\,(V_{ij},\varepsilon_{ij}) \qquad (4)$$

$$\text{with } V_{ij} = V_{ij}(z_{ij}, s_i),$$

where z stands for attributes of the alternatives j and s for attributes of the individuals i.

Let us therefore assume that the parties are not certain how voters will vote when the parties select policy positions in multidimensional policy space. However, all parties have the same subjective expectations about the random behavior of voters. These expectations are represented by probabilistic voting

functions. The probability P_{ij}, that voter i chooses alternative j of the set of alternatives A depends on the observed characteristics z_{ij} of the alternative j compared with the characteristics of each of the other alternatives, as well as on the observed characteristics s_i of the voter. The random utility model specifies this probability as a parametric function of the general form[9]:

$$P_{ij} = f(z_{ij}, z_{ih}, \ \forall j \text{ in } A, \text{and } j \neq h, s_i, \beta) \qquad (5)$$

where f is the function that relates the observed data to the choice probabilities. This function is specified up to some vector of parameters, ß, representing the relative importance of the characteristics. Assume that the random components ε_{ij} are identically and independently distributed with a Gumbel distribution:

$$F(\varepsilon_1,.... \varepsilon_J) = \exp[-\exp(-\varepsilon_1,).....\exp(-\varepsilon_J)] \qquad (6)$$

Given this distribution for the unobserved components of utility, McFadden (1974) derived the conditional multinomial logit model with the following choice probabilities:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{h=1}^{J} \exp(V_{ih})} \quad \forall j, h \in A \qquad (7)$$

Luce (1959) originally derived this model by starting from the axiom of Independence of Irrelevant Alternatives (IIA). The IIA assumption states that the ratio between the choice probabilities of a and b is independent of the set that contains a and b. It therefore implies a form of separability between decisions. The NMNL model is a special case of the Luce model insofar as the representative component V_{ij} is a linear function of the attributes z_{ij}[10].

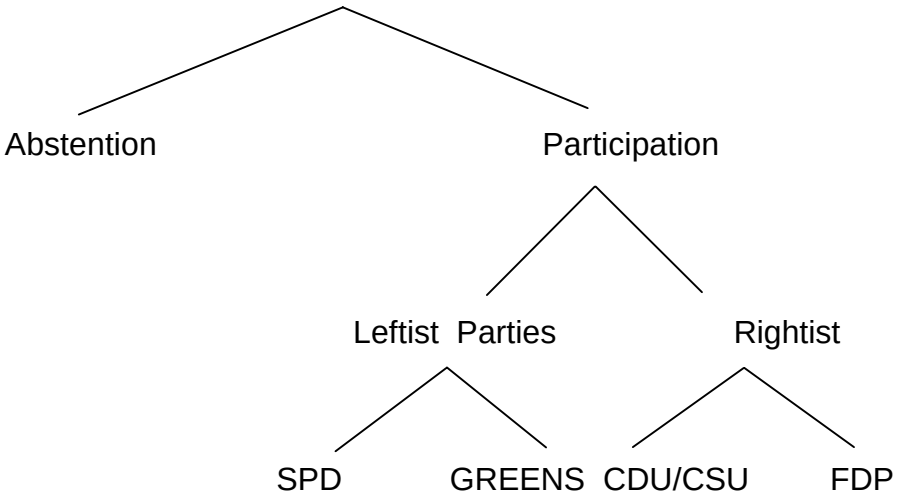
The multinomial logit is inappropriate when IIA holds only for subsets of alternatives but not for all elemental alternatives. Therefore, McFadden (1978, 1981) proposed a more general random utility model which is able to accommodate different degrees of cross-alternative substitution by partitioning of the choice set into nests where IIA holds within nests but not across nests. This so-called generalized extreme value (GEV) model or a special case of it,

the nested multinomial logit model, is a generalization of the multinomial logit model. The general extreme value distribution of the random utility terms is given by:

$$F(\varepsilon_1,\ldots\varepsilon_J) = \exp(-G(\exp(-\varepsilon_1),\ldots,\exp(-\varepsilon_J))), \qquad (8)$$

where G is a nonnegative function of nonnegative elements (y_1,....,y_J), homogenous of degree μ. The nested multinomial logit model allows for the consideration of choice subsets to reflect a nested similarity structure of the alternative in the choice set. For representational aims we represent a three-level tree of a hypothesized decision structure of a simultaneous abstention choice/party choice model (SACPC):

Figure 4: Hypothesized Tree Structure of a Nested Logit Model Applied to SACPC model



where subset A is first split into two subsets, i.e. abstention and participation. Only the participation subset is again split into leftist versus rightist parties, each containing the elemental party alternatives. The multi-stage extension of the extreme value distribution of type B yields the nested multinomial logit model. The joint cumulative distribution of the random variables for a three level tree is given by:

$$G = \sum_{l \in A^*} a_l \left(\sum_{m \in A_l^*} a_{ml} \left| \sum_{j \in A_{ml}^*} \varepsilon_j^{\frac{1}{\mu_{ml}}} \right|^{\frac{\mu_{ml}}{\mu_l}} \right)^{\mu_l}, \qquad (9)$$

$0 < \mu_{ml} \leq \mu_l \leq 1,\ a_{ml}, a_l < 0$,

where A^*, A_l^*, A_{ml}^* denote the choice subsets at the three nesting levels, J is the overall number of alternatives in the choice set A, and the parameters μ are a measure of the correlation of unobserved utility within subsets. Choice probabilities can be derived as follows[11]:

$$P_{lmj} = P_{l \in A^*} \cdot P_{m \in A_l^*} \cdot P_{j \in A_{lm}^*} \qquad (10)$$

where

$$P_{l \in A^*} = \frac{\exp(\alpha_l + \mu_l I_l)}{\sum_{n \in A^*} \exp(\alpha_n + \mu_n I_n)} \qquad (11)$$

$$P_{m \in A_l^*} = \frac{\exp\left| \frac{\alpha_{lm}}{\mu_l} + \frac{\mu_{lm}}{\mu_l} I_{lm} \right|}{\sum_{p \in A_l^*} \exp\left(\frac{\alpha_{lp}}{\mu_l} + \frac{\mu_{lp}}{\mu_l} I_{lp} \right)} \qquad (12)$$

$$P_{j \in A_{lm}^*} = \frac{\exp\left| \frac{\alpha_{lmj}}{\mu_{lm}} + \frac{\beta}{\mu_{lm}} z_{lmj} + \frac{\gamma_{lmj}}{\mu_{lm}} s \right|}{\sum_{q \in A_{lm}^*} \exp\left(\frac{\alpha_{lmq}}{\mu_{lm}} + \frac{\beta}{\mu_{lm}} z_{lmq} + \frac{\gamma_{lmq}}{\mu_{lm}} s \right)} \ . \qquad (13)$$

Variable I, termed inclusive value, summarizes the attributes of the lower level. The presence of this value shows that the variables influencing the party choice probabilities $P_{j \in A^*_{lm}}$ may affect the participation/abstention probability $P_{l \in A^*}$ at the same time. Again: this conceptualization of voting as a multi-level choice process does not imply that the choice at the higher level has to precede temporarily the choice at the lower level. The inclusive values (I_l and I_{lm}) represent the deterministic part of the maximum utility achievable from choosing subsets A^*_l or A^*_{lm}:

$$I_l = \ln \sum_{m \in A^*_l} \exp\left| \frac{\alpha_{lm}}{\mu_l} + \frac{\mu_{lm}}{\mu_l} I_{lm} \right| \qquad (14)$$

$$I_{lm} = \ln \sum_{j \in A^*_{lm}} \exp\left| \frac{\alpha_{lmj}}{\mu_{lm}} + \frac{\beta}{\mu_{lm}} z_{lmj} + \frac{\gamma_{lmj}}{\mu_{lm}} s \right| \qquad (15)$$

Dissimilarity parameters μ indicate the degree of similarity of grouped alternatives in one subset. For the nested logit model to be globally compatible with utility maximization, the estimated values of μ must satisfy the constraint: $0 < \mu_{ml} \leq \mu_l \leq 1$, $a_{ml}, a_l < 0$ (McFadden 1978, Börsch-Supan 1990). If all dissimilarity parameters take values equal to one, and extreme value distribution of type one is implied

$$G = \left| \sum_{j=1}^{J} \varepsilon^{1/\mu} \right|^{\mu}, \ \mu \equiv 1, \qquad (16)$$

the nested multinomial logit model collapses to the simple multinomial logit model.

The estimation procedure, which will be applied in the following, is full information maximum likelihood (FIML)[12] estimation of the NMNL. The loglikelihood function

$$L = \sum_{i} \sum_{l \in A^{*}} \sum_{m \in A_i^{*}} \sum_{j \in A_{ml}^{*}} d_{ilmj} \ln P_{ilmj} , \qquad (17)$$

where index i denotes individuals in the choice set and d_{ilmj} is an indicator variable that takes the value of one if the alternative j has been chosen by individual i and zero otherwise.

The Simultaneous Choice Model of Abstention and Party Vote

Let Z be the K-dimensional policy space. Each individual voter, denoted by i, is assumed to have a well-defined utility function over this policy space. His policy preferences are characterized by a finite point of maximum utility $x_i \in Z$, called his ideal point. Let the voters' choice set be $A = \{a_1,a_2,...,a_j,...,a_J\}$ including the alternative abstention. Voters are assumed to have identical choice sets. Parties are conceived as unitary actors. They take policy positions $z_{jk} \in Z$ in the policy space, where dimensions are assumed to be separable (cf. Ordeshook, 1986: 90). The partial utility derived from a perceived party's policy position z_{ijk} is denoted as $u_i(z_{ijk})$. Each voter chooses the platform closest to his most-preferred positions[13]. Preferences over parties are single peaked and symmetrically distributed around the ideal point. The voter's preference rule can therefore be represented by a distance function: the closer party j's position to x_i, the more probably it is preferred by voter i. We presume the following weighted distance rule for the determination of the total utility function:

$$U_j(z_{ij}) = \sum_{k=1}^{K} -\beta_k \left| x_{ik} - z_{ijk} \right| . \qquad (1)$$

where β is a constant indicating the saliency of each of the K dimensions, which is identical for all i. The saliency weights indicate, how voters value a

single policy dimension, whereas the ratio of weights indicate how they trade off distances in different policy dimensions.

At the same time, non-spatial biases towards parties which are not related to actual policy debates should be considered as well (c.f. Davis/Hinich/Ordeshook, 1970, Enelow/Hinich, 1984). So, Shepsle e.g. points to the fact that, "some bases of voter evaluation are fixed and durable (party 'types are analogous to location-specific capital) while others may be varied by parties (specific issue positions are analogous to product prices). Thus, though there is no direct analogy in politics to prices, the general notion that some features of the firm and its products are durable while others are variable has application to electoral situations" (Shepsle 1991: 43). In order to clarify the implications of a voter's bias towards a specific party we will modify an example provided by Mueller (1989). Let b_{ij} be the bias term representing the propensity of voter i towards party j, with the propensity $b_{ij} > 0$. Choice probabilities in a two-candidate setting can then be presented as follows (see Mueller 1989: 203):

$$\begin{aligned} (P_{i1} > 0{,}5) &\Leftrightarrow (U_{i1} > U_{i2} - b_i) \\ (P_{i1} = 0{,}5) &\Leftrightarrow (U_{i1} = U_{i2} - b_i) \qquad (2) \\ (P_{i2} > 0{,}5) &\Leftrightarrow (U_{i1} < U_{i2} - b_i)\ . \end{aligned}$$

The utility this voter expects from candidate 2's platform must exceed that expected from 1's platform by more than b_{i1}, before 1 loses this individual's vote to 2. This bias term can be represented by the concept of a party-specific loyalty which in the context of rational choice approaches captures the effect of an individual's past voting behavior on the present and future votes and the carry-over effects of past campaigning and party reputation (Fiorina, 1981). Fiorina's 'running tally' model conceives the development of party identification as the ongoing result of the comparative evaluation of the party platforms and performance of party politicians contributing towards a cumulative evaluation of the parties by voters[14]. Then, considering only explanatory variables for the evaluation of the parties $j \in A \backslash a_{abstention}$ leads to the following parametric function:

$$P_{ij} = f(z_{ij}, z_{ih}, b_{ij}, b_{ih} \quad \forall j \text{ in } A \setminus a_{abst.}, \text{and } j \neq h, \beta) \qquad (3)$$

Not all voters participate, however. Therefore, parties are uncertain whether an individual voter will decide to cast a vote at all. Explanatory variables for determining the choice probabilities of abstention versus participation are the already identified policy-related effects of alienation and indifference, which will be modeled as policy-specific in an innovative manner. The smaller the perceived differential between utilities provided by parties' locations the greater the probability of abstention because "the voter feels there just isn't a big enough difference between the candidates to make voting worthwhile" (Enelow/Hinich, 1984: 464). In an extension of the calculus of voting to multicandidate settings McKelvey/Ordeshook (McKelvey/Ordeshook, 1972) criticized that in these contexts the application of B is unclear: "Are benefits greater when a citizen strongly prefers candidate 1 to candidate 2 and 3, and perceives no difference between 2 and 3, or if he perceives some differences among all three" (McKelvey/Ordeshook, 1972: 34). In order to preclude any confusion we try to give a more precise operational definition: The probability of abstaining by indifference increases if $\left|U_i(z_{ijk,min}) - U_i(z_{ijk,min\,2}\right| = 0,$ where $z_{ijk,min}$ represents the perceived location of the minimal-distant party in the k-th dimension and $z_{ijk,min2}$ is the perceived location of the second-distant party in the k-th dimension. When the voter perceives a tie between any of the parties with the one nearest to his ideal point, the probability of abstaining increases. This pragmatic definition of indifference implies that it must be conceived as an expressive component in multiparty systems. It would be conceived as a rational component only when all parties are perceived by voters as taking the same position.

Voters tend, on the other hand, to abstain for reasons of alienation, when the distance towards the minimal-distant party in the k-th dimension exceeds a certain threshold t: "in other words, even a favorite candidate leaves the voter cold" (Enelow/Hinich, 1984: 464). The probability of alienation increases when $\left|x_{ik} - z_{ijk,min}\right| > t$.

Following Riker/Ordeshook, party loyalty is the most important argument of the D term. E contrario, we hypothesize, that being a member of the segment of non-loyals increases the probability of abstaining. Considering only explanatory variables for the decision as to abstaining versus participating, the choice probability takes the form:

$$P(a_{abstention}) = f(d_k, g_k, \bar{b}_j, \varphi), \text{ with } a_{abstention} \in A. \quad (4)$$

where d_k stands for a policy-specific alienation, g_k represents a policy-specific indifference, $\bar{b}_j$ indicates that the voter has no party-specific loyalty, and φ is the vector of respective weights. An empirical application of this model will now be provided.

Data

The following empirical analyses are based on the German part of the international 'Comparative National Election Project' (CNEP)[15] containing the national study of the first all-German general election of December 2, 1990. For simplification, the following analyses will be restricted to the first wave of the West German study[16] with a target sample size of N=1400.

In this secondary analysis we have chosen several position issues measured by bipolar seven-point policy scales. These 'constructed subjective attribute scales' (Keeney/Raiffa 1992: 40) have been introduced in the US NES since 1964 and are also contained in the CNEP study. They have been developed in order to evaluate public attitudes toward public policies. The data set contains four subjective attribute scales for the higher-level objectives: German unification, immigration policy, abortion, and nuclear energy. Policy-specific distances has been calculated for the four major German Parties CDU/CSU, SPD, FDP, and the GREENS and specified as generic variables[17].

Long standing party biases will be captured by the concept of party loyalty as measured by a version of the classic party identification question which explicitly accentuates its long term aspect.

The voting choice is operationalized by the respondent's prospective report of his or her vote for one of the parties resulting in a multicategorial dependent

variable. The German electoral rules allow the voter to cast two votes. The CNEP-study, therefore, differentiates between the first and the second vote. Since the share of seats in the federal parliament, the Bundestag, depends on the second vote, where the voter indicates his preference for a party, the second vote will be predicted in the empirical analysis. Due to the very limited number of minor party responses, the following analysis is carried out for the major four West German alternatives: SPD, CDU/CSU, FDP, and the Greens[18]. Due to the 5%-threshold of the German proportional system, the Greens as the West German part of the ecologists, has not been represented in the German Bundestag (1990-1994).

The problem of the analysis carried out here is that the proportion of non-voters is principally underrepresented in surveys, due to effects of social (un)desirability. Therefore, an indicator of abstention is used, recording not only those who reveal their abstention, but also potential-non-voters who indicate that they are not sure of going to the ballot. Due to an extremely small response category abstention in the East German sample, a corresponding analysis could not be carried out.

Estimation Results

To the surprise of most analysts, the first all-German election to the German Bundestag in November 1990 led to the lowest turnout in the post-war period in West Germany: only 78 % of the electorate participated. In the East German part, turnout dropped from 93 % in the election to the 'Volkskammer' in May 1990 to 74 % in the general election. Explanations range from the assumed growing of a general reluctance towards parties ('Parteiverdrossenheit') in West Germany, to doubts and ambivalence concerning the political realization of the unification, up to the assumption that the conviction of victory of government supporters on the one hand and the fatalism of the supporters of the opposition on the other were the reason for the small participation rate.

Before estimating a simultaneous model of abstention choice and party choice, we will first follow the implicit rule in previous analyses, namely to structure the decision as two separate decisions. Table 1 shows results for a model which explains the choice of the four parties, indicated below, by policy distances and

a bias term as suggested by the now classical model of Davis/Hinich/Ordeshook (1970). The estimation results for this model reveal that party loyalty, measuring a voter's bias towards a party and specified as a generic variable, has an outstanding effect on the party choice also in multiparty systems. This effect is identical for all parties[19]. Nevertheless, policies specified as generic distances[20] have an impact as well. We consider this as the most striking message of this table. All issues have negative signs: i.e. the higher the perceived distance to a party in a given policy, the lower the probability of this party to be voted for.

Table 1: Policy Distances, Party-specific Loyalty and Party Choice /West Germany

Variable	**ß**	**t-Ratio**	**P-Value**
Party Loyalty	2.088	16.177	0.000
Unification	-0.142	-2.352	0.019
Immigration	-0.245	-5.436	0.000
Abortion	-0.050	-1.203	0.229
Nuclear Energy	-0.312	-6.183	0.000
CDU/CSU-Constant	0.702	4.203	0.000
SPD-Constant	0.339	1.963	0.049
GREENS-Constant	0.081	0.395	0.693
2*(LL(N)-LL(0)) = 870.043, DoF = 5, P-Value = 0.000, Pseudo R^2 = 46.9 %			
Percentage of Correct Classifications: 80.5 %			

(N =775)

The ranking of saliency weights is surprising, however: In the case of the West German population, nuclear energy and immigration are more important issues than German unification. This pattern is in contrast to the results for East Germany (cf. Thurner, 1998) where unification stood up as the most important issue. The reason for this pattern may be peculiarities of the West German agenda development[21]: Immigration policy was a highly politicized issue since the mid-1980s and ranked always as one of the most important political problems. This applies also to nuclear energy as an indicator for ecological issues in general, which has been chosen not only by the GREENS but also by the SPD as one of the central campaigning issues before unification intervened unexpectedly. For many West Germans this new and very difficult issue seems to have been of minor importance or fully absorbed by long standing biases.

In order to test whether cross alternative substitution of this MNL model is equal for all parties, we applied the Hausman/McFadden-test[22] to various subsets of the four-alternative set underlying our MNL model. These tests indicate that the multinomial model does not violate the assumption of independence of irrelevant alternatives[23]. So the specification of an MNL seems to be adequate. Goodness-of-fit indices indicate a relatively high explanatory power of the variables.

In the next model, the binary choice between abstention (1) and participation (0) is predicted by policy-specific alienation and indifference as well as the self-reported non-loyalty to each of the parties:

Table 2: Policy-Specific Alienation and Indifference, Party Loyalty and Abstention /West Germany

Variable	ß	t-Ratio	P-Value
No Party Loyalty	1.810	9.192	0.000
ALIENATION			
Unification	0.007	0.048	0.961
Immigration	0.186	1.938	0.053
Abortion	0.149	0.976	0.329
Nuclear Energy	0.140	0.871	0.384
INDIFFERENCE			
Unification	-0.267	-1.358	0.175
Immigration	0.237	1.389	0.165
Abortion	-0.211	-0.994	0.320
Nuclear Energy	-0.181	-0.882	0.378
Constant	-2.294	-10.494	0.000
2*(LL(N)-LL(0)) = 99.188, DoF = 9, P-Value = 0.000, Pseudo R^2 = 12.2 %			
Percentage of Correct Classifications: 84.1 %:			

(N = 923)

As has been expected, the propensity of members of the non-loyal segment to abstain is higher than in the party loyal segment. When applying a significance level of 5 % this is the only significant effect. Perceived alienation in the immigration issue is the only variable in the indifference block, whose coefficient is significantly larger than zero at the 10% level. Signs are in the expected directions with the exception of unification, abortion, and nuclear energy in the indifference block. As a conclusion it can be stated that long-standing trust and loyalty towards a party is an important factor for explaining the participation choice in this specific election. In the argumentational terms of Brody/Page (1973) and Guttmann/Hilger/Shachmurove (1994) this would point to a preponderance of expressive effects. However, before coming to a final conclusion, we first present a systematic overview of the results of different SACPC models:

Table 3: Overview of Tested SACPC Models

	Model	Dissimilarity Parameter(s)	$2(LL_{NMNL}-LL_{MNL})$
Tree I (figure 1):	not specified	-	-
Tree II (figure 2):	NMNL	particip.: 1.614, s.e.: 0.255	$\chi 2 = 1.732$, DoF=1
Tree III (figure 3):	MNL	-	-
Tree IV (figure 4):	NMNL, level 1	leftist: 1.640, s.e.: 0.171	$\chi 2 = 31.962$, DoF=3
	level 1	rightist: 1.654, s.e.: 0.216	
	level 2	particip.: 1.198, s.e. 0.201	

We have not been able to model tree I, as we do not have data on the (multidimensional) alternative of abstention after having chosen a party already. This is due to the fact that our data set reports only observed final choices.

We have applied three different specification tests to compare the model specifications corresponding to trees II, III, and IV. Focusing upon party choice only, we find inconclusive results: Hausman-McFadden's specification test indicates that parties should not be further aggregated into subsets of different political orientation. A likelihood ratio test comparing tree structures II and IV, however, leads to the contrary result[24]. The estimates for the lower-level dissimilarity parameters of tree IV are significantly larger than one, rejecting the MNL model and leaving doubts as to whether this NMNL model is compatible with stochastic utility maximization[25]. However, estimation results for the models corresponding to trees II, III, and IV reflect a largely identical pattern of impact for all explanatory variables. We therefore decided not to distinguish subsets of parties in the following.

Regarding the decision either to choose one of four parties or to abstain, both the t-ratio for the upper-level dissimilarity parameter for tree IV and the likelihood ratio test comparing models II and III indicate that abstention choice and party choice should not be modeled as separate subsets of an NMNL model for the given set of explanatory variables. In other words, the decision to participate and to choose a party can be considered as non-hierarchical. We therefore present estimation results for tree III only and use these in order to test our theoretical model[26].

Table 4: SACPC, Tree III /West Germany

Variable	ß	t-Ratio	P-Value
PARTY CHOICE			
Party Loyalty	2.287	17.411	0.000
Unification	-0.111	-1.933	0.053
Immigration	-0.224	-5.159	0.000
Abortion	-0.306	-0.803	0.422
Nuclear Energy	-0.274	-5.732	0.000
ABSTENTION			
No Loyalty	0.432	1.976	0.048
ALIENATION			
Unification	-0.117	-0.809	0.419
Immigration	-0.004	-0.042	0.966
Abortion	0.082	0.526	0.599
Nuclear Energy	-0.095	-0.572	0.567
INDIFFERENCE			
Unification	-0.165	-0.819	0.413
Immigration	0.421	2.094	0.036
Abortion	-0.224	-1044	0.296
Nuclear Energy	0.031	0.150	0.881
CDU/CSU-Constant	0.605	3.689	0.000
SPD-Constant	0.251	1.477	0.139
GREENS-Constant	0.048	0.246	0.806
Abst.-Constant	-0.229	-0.787	0.431
2*(LL(N)-LL(0)) = 1004.770, DoF=14, P-Value = 0.000, Pseudo R^2 = 37.6 %			
Percentage of Correct Classifications: 56.1 %			

(N = 923)

Table 4 shows that patterns in the block of variables explaining party choice remain unchanged compared to the model in table 1 and will, therefore, not be commented further. When simultaneously explaining abstention and party choice, the expressed non-loyalty towards each of the parties remains the most important criterion for abstention, but its influence is smaller as compared to the separate decision model (table 2). It remains to be tested in further studies, if this applies also to other settings. The central result in this simultaneous MNL model is that, contrary to the binary model of participation choice, the only

significant spatial effect is now the perceived indifference in the immigration issue. Controlling for comparative party distances in different policy dimensions implies that the effect of alienation in the immigration issue, which has been only weakly significant in table 2, completely disappears in table 4. Now, the last model reveals that a perceived indifference in this important issue significantly augments the probability of abstaining. Regarding contents it may therefore be argued, that when controlling for all theoretically important factors, only the lack of different party positions in the immigration issue leads to an increase of the abstention probability. Increases in perceived indifference in highly salient issues results in an increase of the probability of abstaining.

Conclusion

This study addresses McKelvey's question: "With what probability will citizens with ideal points at x vote for one candidate, the other, or abstain?" (McKelvey, 1975). Our approach is empirical, using data on the German general election in 1990, and refers to a spatial voting model that integrates the choice of parties and abstention. From a methodological point of view, our analysis differs from previous empirical analyses, in which abstention choice and party/candidate, choice have always been modeled separately. In the context of nested logit models we discussed the meaning of sequentiality and simultaneity of decisions. We derived several possible decision trees and tested their adequacy to represent the voter's choice.

From a substantive view, this paper shows, first, that a hierarchical nesting structure is not necessary. The simultaneous abstention and party choice can be represented by a (non-hierarchical) multinomial logit model or by two separate logit models. The impact structure of the explanatory variables remains the same when comparing the estimation results for separate models and the simultaneous model. Non-spatial, party-specific biases prove to be the most important explanatory variables, but political issues do matter. Perceived indifference in the immigration issue is the only policy-related factor leading to an increased propensity of non-voting.

Notes

1. For an overview see Enelow/Hinich 1984, Enelow/Hinich 1990.

2. See also Davis/Hinich/Ordeshook 1970. An extension to multi-candidate contests is McKelvey/Ordeshook 1972. This article focuses especially on the aspect of closeness of election outcomes and is therefore not taken into consideration in the following.

3. For an overview of the vaste literature see Mueller 1989, Struthers/Young 1989, Brennan/Lomasky 1993, Grofman 1995, Uhlaner 1995.

4 In the following we do not consider decisions under risk, where the concept of decision trees has been intensely discussed, see Raiffa 1968.

5. CDU/CSU: Christian Democrats, SPD: Social Democrats, FDP: Liberal Democratic Party, GREENS: Ecologists.

6. "...one must distinguish between hierarchical behavior and a hierarchical structure for the mathematical forms of the choice probabilities" (Pudney, 1989: 125).

7. See also Hensher/Johnson, 1981, Ben-Akiva/Lerman, 1985.

8. For the explicit introduction of random utility models to represent binary candidate contests, see especially Coughlin, 1992.

9. Cf. Train, 1986: 8.

10. The underlying idea of this model is the view of commodities as bundles of properties that make up their utility, earlier advocated by Lancaster (1966) as a general approach to consumer demand.

11. The individual-specific index i will be neglected for reasons of simplification.

12. Models have been estimated with LIMDEP 7.0.

13. We preclude, therefore, the possibility of ‘strategic voting’.

14. Marketing researchers generally introduce a measure of brand loyalty into the utility function to account for differences in utility across households and over time. One of the most widely used measures of brand loyalty, proposed by Guadagni and Little (Guadagni/Little, 1983), is an exponential smoothing model of past choice behavior by the household. As we cannot observe long-term individual vote histories, we content ourselves with achieving a cross-sectional preference segmentation by the concept of party loyalty.

15. The German part of the project has been financed by Deutsche Forschungsgemeinschaft.

16. This is due to differently structured party systems in East and West Germany. A detailed comparative analysis of voting decisions in East and West Germany can be found in Thurner (1998).

17. "The characteristics of a generic variable are that it varies in value across all response categories and has an associated generic parameter which remains constant in value across response categories" (Wrigley, 1985: 75), for further information on types of variables see Ben-Akiva/Lerman (1985), Thurner (1998).

18. As criterion for inclusion into the choice set, at least 30 cases per alternative should remain.

19. The respective likelihood ratio test shows that the splitting of the generic variable into party-specific biases is not adequate: χ^2 =4.83, DoF = 3.

20. Parameters of policy distances are identical for all parties. Unification: χ^2 =1.35, DoF = 3; immigration: χ^2 = 7.65, DoF= 3; abortion: χ^2 =1.24, DoF = 3; nuclear energy: χ^2 =6.46, DoF = 3.

21. For more details see the respective time series of stated 'most important political problem' in Thurner, 1998.

22. The form of the test statistic proposed by Hausman/McFadden (1984) is: $\chi^2(c) \sim (\hat{\beta}^r - \hat{\beta}^u)' \left[\hat{C}ov(\hat{\beta}^r) - \hat{C}ov(\hat{\beta}^u)\right]^{-1} (\hat{\beta}^r - \hat{\beta}^u)$, where: $\hat{\beta}^u$: estimator based on full choice set, $\hat{\beta}^r$: estimator based on restriced choice set, $\hat{C}ov(\hat{\beta})$: estimated covariance matrix of estimator

23. Removing CDU/CSU yields: c=9.423, DoF=7, P-value=0.223; SPD: $c \approx 0$, DoF=7, P-value ≈ 1;FDP: c=11.751, DoF=7, P-value=0.109; GREENS: c= 8.873, DoF=7, P-value= 0.261.

24. Contradictory results of the three specification tests are not uncommon for nested multinomial logit models (Cameron, 1985, Eymann 1995) when information on relevant explanatory variables is partly unavailable.

25. Cf. Börsch-Supan, 1990.

26. d IV Estimation results corresponding to trees II anwill be sent to interested readers on request.

Appendix

Issue Scales:

German Unification:

"The economic rehabilitation of the former GDR is regarded as the most urgent problem of German reunification. There are different opinions regarding the best policies for this rehabilitation. Some argue that it would be best if the state interferes as little as possible. Others argue that a far-reaching state intervention would be the best. What is your opinion? Use the 1-7 scale to indicate your own position.

1 represents few state interventions, 7 far-reaching state interventions. You can differentiate with the values between."

Immigration:

"The German Constitution grants asylum to victims of political persecution from other countries. The discussion at the moment is if this right of political asylum should be restricted or not. What is your opinion?"

Abortion:

"There is a lot of discussion about the state's position towards abortion. Here there are also quite different points of view. One group says that the state should generally make abortion a punishable offense, the other group says that the woman herself should decide about possible abortion. What is your opinion?"

Nuclear Energy:

"Concerning nuclear energy, there are as well contradictory attitudes. Some say the use of nuclear energy should be further extended, others say all nuclear power plants should be switched off immediately. What is your opinion?"

Party Loyalty:

"Many people lean towards one particular party over a long period of time, although they do vote for another party once in a while. How about you: Do you - generally speaking - lean towards a particular party and if yes, to which one?"

"How strong is -over all - your affiliation to this party: very strong, rather strong, moderate, rather weak, very weak?"

Voting Choice (First and Second Vote):

"The next national parliamentary election will take place on December, 2, this year as an all-German election. Will you go to the polls?....At this national election you have two votes: the first vote for electing the constituency representative, the second vote for electing the 'Länder' list of a party. Here is

a ballot card similar to the one you will get in the national election. The candidate of which party would you vote for?...And which party will you vote for with your second vote?

Participation Choice:

"The next general election will take place on December, 2 this year as an all-German election. Will you go to the polls?: Yes, No, Don't Know".

References

Anderson, S. P. and G. Glomm, 1992: Alienation, indifference and the choice of ideological position. Social Choice and Welfare 9: 17-31.

Ben-Akiva, M. and S. R. Lerman, 1985: Discrete choice analysis. Cambridge (Ma.): MIT Press.

Börsch-Supan, A., 1987: Econometric analysis of discrete choice with applications on the demand for housing in the US and West Germany. Berlin, New York: Springer-Verlag.

Börsch-Supan, A., 1990: On the compatibility of nested logit models with utility maximization. Journal of Econometrics 43: 371-387.

Brennan, G. and L. Lomasky, 1993: Democracy and decision: The pure theory of electoral preference. New York: Cambridge University Press.

Brody, R. A. and B. I. Page, 1973: Indifference, alienation, and rational decisions: The effects of candidate evaluation on turnout and the vote. Public Choice 15: 1-17.

Cameron, T. A., 1985: A nested logit model of energy conservation activity by owners of existing single family dwellings. The Review of Economics and Statistics 67: 205-221.

Coughlin, P., 1992: Probabilistic voting theory. Cambridge: Cambridge University Press.

Dalton, R. J., 1993: The new Germany votes. Unification and the creation of the new German party system. Oxford: Berg Publishers.

Davis, O., M. J. Hinich and P. C. Ordeshook, 1970: An expository development of a mathematical model of the electoral process. American Political Science Review 64: 426-448.

Downs, A., 1957: An economic theory of democracy. New York: Harper & Row.

Coughlin, P. 1992: Probabilistic voting theory. Cambridge: Cambridge University Press.

Enelow, J. M. and M. J. Hinich, 1984: The spatial theory of voting. Cambridge: Cambridge University Press.

Enelow, J. M. and M. J. Hinich, 1990: Advances in the spatial theory of voting. Cambridge: Cambridge University Press.

Eymann, A., 1995: Consumer's spatial choice behavior. Heidelberg: Physica-Verlag.

Fiorina, M. P., 1981: Retrospective voting in American national elections. New Haven: Yale University Press.

Grofman, B., 1995: Is turnout the paradox that ate rational choice theory? In B. Grofman (Ed.), Information, Participation, and Choice, 93-103. Ann Arbor: The University of Michigan Press.

Guadagni, P and J. D. C. Little, 1983: A logit model of brand choice calibrated on scanner data. Marketing Science 2: 203-238.

Guttman, J. M., N. Hilger and Y. Shachmurove, 1994: Voting as investment vs. voting as consumption: New evidence. Kyklos 47: 197-207.

Hausman, J. and D. McFadden, 1984: Specification tests for the multinomial logit model. Econometrica 52: 1219-1240.

Hensher, D. A. and L. W. Johnson, 1981: Applied discrete choice modeling. New York: Halsted Press.

Hinich, M. J., J. O. Ledyard and P. C. Ordeshook, 1972: Nonvoting and the existence of equilibrium under majority rule. Journal of economic theory 14: 144-153.

Keeney, R. L. and H. Raiffa, 1992: Decisions with multiple objectives: Preferences and value tradeoffs. Cambridge: Cambridge University Press (first published: New York: John Wiley & Sons, 1976).

Kirchgässner, G., 1992: Towards a theory of low-cost decisions. European Journal of Political Economy 8: 305-320.

Lancaster, K. J., 1966: A new approach to consumer theory. Journal of Political Economy 74: 123-157.

Manski, C. F., 1977: The structure of random utility models. Theory and Decision 8: 229-254.

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

McFadden, D., 1978: Modeling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars and J. W. Weibull (Eds.): Spatial interaction theory and planning models, 75-96. Amsterdam: North-Holland.

McFadden, D., 1981: Econometric models of probabilistic choice. In Charles Manski and Daniel McFadden (Eds.): Structural analysis of discrete data with econometric applications. Cambridge, Mass.: MIT Press.

McKelvey, R., P. C. Ordeshook, 1972: A general theory of the calculus of voting. In J. F. Herndon (Ed.): Mathematical applications in political science, Vol. 5, 32-78. Charlottesville: University of Virginia.

McKelvey, R., 1975: Policy related voting and electoral equilibrium. Econometrica 43: 815-843.

Mueller, D., 1989: Public choice II. Cambridge: Cambridge University Press.

Pudney, S., 1989: Modelling individual choice. The econometrics of corners, kinks and, holes. London: Basil Blackwell.

Raiffa, H., 1968: Decision analysis. Reading, Massachusetts: Addison-Wesley.

Riker, W. H. and P. C. Ordeshook, 1968: A theory of the calculus of voting. American Political Science Review 62: 25-42.

Riker, W. H. and P. C. Ordeshook, 1973: An introduction to positive political theory. Englewood Cliffs: Prentice Hall.

Shepsle, K. A., 1991: Models of multiparty electoral competition. Chur: Harwood Academic Publishers.

Stigler, G. J., 1972: Economic competition and political competition. Public Choice 13: 91-106.

Thurner, P. W., 1998: Wählen als rationale Entscheidung. Die Modellierung von Politikreaktionen im Mehrparteiensystem unter Anwendung von multiattributiven Zufallsnutzenmodellen. München: Oldenbourg, Scientia Nova (in print).

Smithies, A., 1941: Optimum location in spatial competition. Journal of Political Economy XLIX: 423-439.

Struthers, J. and A. Young, 1989: Economics of voting: Theories and evidence. Journal of Economic Studies 16: 1-42.

Train, K. E., 1986: Qualitative choice analysis: Theory, econometrics, and an application to automobile demand. Cambridge (Ma.): MIT Press.

Tversky, A., 1972: Elimination by aspects: A theory of choice. Psychological Review 79: 281-299.

Tversky, A., 1972: Choice by elimination. Journal of Mathematical Psychology 9: 341-367.

Tversky. A., A. Sattah, 1979: Preference trees. Psychology Review 86: 542-573.

Uhlaner, C. J., 1995: What the Downsian voter weights: A reassessment of the costs and benefits of action. In Bernard Grofman (Ed.): Information, participation, and choice, 67-79. Ann Arbor: The University of Michigan Press.

Wrigley, N., 1985: Categorical data analysis for geographers and environmental scientists. London: Longman.

Keywords: Spatial theory of voting, paradox of non-voting, multiparty system, nested multinomial logit model, random utility theory