LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
INSTITUT FÜR STATISTIK

Examining Undecided Voters in Multiparty Systems - An Approach with Discrete Choice Models under Ontic Imprecision

Masterthesis

Author	Dominik Kreiß
Supervisor	Prof. Dr. Thomas Augustin
Date	Munich, October 24, 2019

Contents

1 Introduction: Pre-Election Behavior Modeling **5**

2 Methods **7**

2.1 Discrete Choice Models 7

2.1.1 From Random Utility to the Multinomial Logit Model 7

2.1.2 The Multinomial Logit Model as a Discrete Choice Model 8

2.1.3 Estimation in the Multinomial Logit Model 9

2.1.4 Regularization and Variable Selection in the MNL 10

2.2 Ontic and Epistemic Imprecision 12

2.3 Discrete Choice Models under Ontic Imprecision 13

3 Discrete Choice Models for the 2017 German Federal Election under Ontic Imprecision **15**

3.1 The Data from the German Longitudinal Election Study 15

3.1.1 Background and Emergence of the Data 15

3.1.2 Construction of the Variable "Ontic" 16

3.1.3 Pre-Selection and Variable Description 18

3.1.4 Descriptive Statistics 21

3.2 The Model 22

3.2.1 Applied Regularization 23

3.2.2 Model Specification 24

3.3 Results and Interpretation 25

4 Concluding Remarks **30**

List of Figures

1 Frequencies of the degree of certainty attached to the intent within the data 17

2 Number of observations for each constructed ontic category within the data 18

3 Illustration of the item nonresponse concerning the chosen variables in the data . . . 20

4 Visualization of the regularization applied to the binary variables chosen, dependent on the lambda value on the x-axis and the estimate on the y-axis 24

List of Tables

1 Display of the connection between the conventional and ontic numbers of observations for each category 21

2 The mean of the binary variables split up between the ontic categories 22

3 The mean of the binary variables split up between the conventional categories 22

4 Results of the individual specific variables from the ontic model 26

5 Results of the individual specific variables from the conventional model 28

6 Comparison of the results of the individual specific variables from ontic and conventional models for the main six parties for the second vote 28

7 Comparison of the results of the individual specific variables from ontic and conventional models for the main six parties for the first vote 29

Abstract

The rising number of undecided individuals in pre-election polls constitutes a severe problem for conventional surveys and methodology, permitting only precise answers to the question about which party respondents plan to vote for. Hereby, these growing subgroups, undecided between specific parties are neglected and uncertainty attached to the answers is not accounted for. This thesis proposes to extend the traditional question about party preference to the possibility of a set valued answer, if the individual is incapable to state one precise party, allowing undecided participants to express their position in an accurate manner. The resulting set valued information can be analyzed under ontic or so-called epistemic imprecision, while we focus on the ontic, considering each group of individuals undecided between specific parties as an entity of its own and therefore enabling choice modeling for these groups of interest. Hereby, the state space of the original nominal and finite set of the single parties is extended to its power set, on which all conventional methodology is applicable. The new approach of discrete choice models under ontic imprecision is implemented using the dataset from the *German Longitudinal Election Study* for the federal election in 2017, providing sufficient information for an artificial construction of an ontic variable. The multinomial logit model is used for choice modeling and applied to this data using regularization for variable selection. Within the results, some differences to the conventional model are apparent, stressing the virtue of our novel, more differentiated and accurate approach, while more detailed information is provided by including the groups undecided between specific parties. This thesis can be seen as a step towards the solution of the chicken-egg-dilemma resulting from the lack of surveys including the set valued question as well as missing methodology for such data.

1 Introduction: Pre-Election Behavior Modeling

Pre-election polls and surveys about electoral behavior are widely used with great media attention, even though criticized heavily for wrong predictions during the past years. (Jennings and Wlezien, 2018) (Keeter, 2018, p. 5) The centre of these surveys usually is the intended party to vote for. In conventional surveys the participant is allowed to choose one single party (often only out of the ones anticipated to reach at least one seat in the parliament), an *other* option and a *do not know* option, while several other questions are amended, depending on the survey at hand. The analysis of such data mostly aims at one of two goals. One: the prediction of the election outcome and two: the analysis of voting behavior. This thesis will focus on the latter and only touch the former briefly. Frequently, participants choosing the *do not know* or *other* options are excluded in conventional analysis. (Plass et al., 2015, p. 257)

One major development concerning pre-election polls is the rising number of undecided voters. A study argues that the number of undecided individuals eligible to vote one month prior to the federal German election has risen from 39 to 46 % in the last two elections. (Zeit, 2017, p. 1). The increase of undecided voters inevitably leads to more individuals being neglected in the analysis, as long as participants are given the three options explained above and only the precise party choice answers are considered. If the underlying process of being an undecided voter is not *missing completely at random*, next to the decrease of sample size, bias is introduced if this problem is not accounted for. (Rubin, 2004, ch. 1)

In order to live up to this challenge, this thesis proposes to allow undecided voters a fourth option of set valued answers. Hereby, the undecided are not forced to choose solely one single party, but are allowed to state all parties they are incapable to choose from, hence excluding parties they will definitely not vote for. Therefore, the participants can express their opinion no matter how precise or imprecise it might be, while all individuals are compelled to state their intend as precisely as they are capable of. Hereby, the complete set becomes the *do not know* option and the empty set is the information of participants not intending to vote at all, while the set representing exactly one party is the decided voter.

The idea of set valued response in election choice was also recently introduced by (Oscarsson and Rosema, 2019) in a political science framework, arguing that stepwise exclusion of options is the natural process of human choice (ibid., p.256). From this research, four further papers (Oscarsson and Oskarson, 2019), (Rekker and Rosema, 2019), (Fredén and Sohlberg, 2019) and (Steinbrecher and Schoen, 2019) resulted in cooperation, mostly focussing on political implications (ibid.) and the modeling of whether or not people have choice sets (Oscarsson and Oskarson, 2019, p. 278 ff.), but not providing a proper mathematical framework for the analysis of the resulting set valued information.

The expected improvement concerning the nonresponse and accuracy comes at the price of imprecision attached to the set valued answers of the undecided. This so called fuzzy information is a topic thoroughly researched for example in (Couso and Dubois, 2014). The resulting set valued information can be analyzed under ontic or epistemic imprecision, the latter understanding the set as a container of one accurate value, while the former regards the set as an entity of itself. Both approaches, in contrast to the conventional one, enable inclusion of the undecided voters. While the epistemic approach recommends itself for prediction purposes, the ontic one enables analyzing subgroups undecided between specific parties as entities of their own, therefore providing information on groups of special interest concerning voting behavior. This thesis focusses on the ontic approach, regarding the set valued information of an undecided as the most accurate possible and true representation of the opinion at the given point in time. Hereby, the nominal and finite state space containing the single parties used by the conventional analysis is extended to its power set (Couso and Dubois, 2014, p. 1504), enabling usage of all common statistical methods on this, once again finite and nominal new space. (Plass et al., 2015, p. 2)

In this thesis we make use of this fact, extending discrete choice models first introduced by (McFadden, 1973) to ontic imprecision. State of the art choice modeling as can be found under (Mauerer et al., 2015) or (Tutz, 2011, ch. 8) is applied on the new state space following the new, ontic approach.

The idea is implemented using data from the most renowned German pre-election poll, the *German Longitudinal Election Study* (GLES) for the federal election of 2017, which even though not containing set valued response, provides information about the uncertainty of the vote and assessment of the parties, enabling artificial construction of an ontic, set valued variable. On the one hand the implementation strives to prove the concept of choice modeling under ontic imprecision and on the other provides new insights into the German federal election of 2017, by focussing on the specific subgroups. Hereby, we contribute to a solution to the chicken-egg dilemma, resulting from the lack of surveys including the set valued question as well as missing methodology, providing a straight forward, applicable approach for such data. (Fink, 2018, p. 7) The analysis is conducted using the *Categorically Structured Lasso* regularization method from (Tutz, Pößnecker, and Uhlmann, 2015) to select variables for the multinomial logit model as the most common discrete choice model. (Tutz, 2011, ch.8) The results of the ontic model are interpreted and compared to the conventional approach.

Our main contributions are: 1: Providing the mathematical framework for straight forward and applicable methodology for set valued responses in pre-election polls for choice modeling. 2: Introducing and applying discrete choice models under ontic imprecision for pre-election research. 3: Providing insights into behavioral patterns of subgroups undecided between specific parties for the German federal election of 2017.

This thesis is structured as follows: First, the theory and genesis of discrete choice models will be reviewed in chapter 2.1 and after explaining ontic and epistemic imprecision in chapter 2.2, further extended to ontic imprecision in chapter 2.3. Chapter 3 applies the earlier discussed methods to the German federal election in 2017 using data from the GLES. Later on, results will be analyzed, the approaches compared and the success of the concept will be discussed, while providing ideas for further research.

2 Methods

The methods section is divided into three parts. First, the multinomial logit model in the context of discrete choice models are motivated and explained, the estimation process discussed and regularization as a mean of variable selection reviewed. Second, ontic and epistemic imprecision as approaches to deal with set valued imprecision are compared, while providing the basic mathematical principles. Third, discrete choice models are introduced to ontic imprecision, laying the foundation for the application in chapter 3.

2.1 Discrete Choice Models

Due to the human categorical understanding of the world and choice (M. Young, 2016, p. 43 ff.), people are frequently faced with finite and therefore discrete options to choose from. Modeling behavior being faced with these options has a long tradition in social science, economics and statistics. (Thurner, 2000, p. 493) The current research mostly goes back to the works of (McFadden, 1973) but extensions and different approaches have been discussed extensively. An overview of discrete choice models in the realm of political science can be found for example in (Thurner, 2000).
This chapter motivates the multinomial logit model, as the most frequently used discrete choice model with the random utility assumption. Hereby, the notation and motivation is provided in chapter 2.1.1 following the approach of (Tutz, 2011, ch. 8) as well as the estimation in chapter 2.1.3. Further on, regularization as a mean of variable selection in discrete choice models is discussed.

2.1.1 From Random Utility to the Multinomial Logit Model

This section provides a probabilistic modeling approach analyzing behavior of individuals facing discrete and disjunctive options. Our goal is to model the probability $P_{i,B}(r)$ for individual i to choose option r from the set of alternatives $B = 1,...,k$ based on empirical data about the individual, the options and the choices made. Hereby, the interpretation of influences on the individuals' choice is the main interest. Basis for such a probabilistic choice model for $P_B(r)$ is a *probabilistic choice system*: $\{$B, $\{P_B$, B $\in B\}\}$ with B as the set of available alternatives and $P_B(r)$ as the probability to choose option r from the set B. Given this *probabilistic choice system* we assume that each option $r \in B$ holds a latent utility U_r divided in an observable part u_r and a random part ε_r. In our case of electoral choice the fixed utility u_r is characterized by attributes of the alternatives and individual characteristics.
Second, we open the utility maximization assumption, arguing that each individual always chooses the option for which the utility is maximal, hence

$$Y = r \leftrightarrow U_r = \max_{1,...,k} U_j \tag{2.1}$$

If a probabilistic choice system meets these assumptions, it is called a *random utility model*.

The probability $P_B(r)$ to choose option r from the set B can therefore be constructed as

$$\begin{aligned} P(U_r \mid B) &= P(U_r \geq U_k,\ \forall k \in B) \\ \Leftrightarrow P(u_r - u_k &\geq \varepsilon_k - \varepsilon_r,\ \forall k \in B) \end{aligned} \tag{2.2}$$

This equation leads to the distribution function $F_{B,r}$ over the differences $\varepsilon_{sr} = \varepsilon_s - \varepsilon_r$ as a $m-1$ dimensional vector $\varepsilon_{B,r} = (\varepsilon_{1r}, ..., \varepsilon_{mr})$ under the iid assumption:

$$F_{B,r} = \int_{-\infty}^{u_r - u_1} \cdots \int_{-\infty}^{u_r - u_m} f_{B,r}(\varepsilon_{1r}, \ldots, \varepsilon_{mr}) \quad d\varepsilon_{1r}, \ldots, d\varepsilon_{mr} \tag{2.3}$$

with $f_{B,r}$ as the density of ε_r and $F_{B,r}$ as the cumulative distribution function of $\varepsilon_{B,r}$. Now for every choice set B of which the elements of B consist of the entire discrete alternatives, any distribution assumption about $\varepsilon^T = (\varepsilon_1, ..., \varepsilon_k)$ will result in a specific type of a discrete choice model.

Choosing the Gumbel distribution $F(x) = \exp(-\exp(-x))$ of independent and identically distributed $(\varepsilon_1, ..., \varepsilon_k)$ results in the cumulative distribution function of the differences $(\varepsilon_1 - \varepsilon_r, ..., \varepsilon_k - \varepsilon_r)$ in regards to the $m-1$ number of x_i from the Gumbel distribution following (Tutz, 2011, p. 227):

$$F_{m-1}(x_1, ..., x_{m-1}) = \frac{1}{1 + \sum_{i=1}^{m-1} \exp(x_i)} \tag{2.4}$$

This leads directly to the so called multinomial logit model (MNL) to calculate the probability of one option $r \in B$.

$$P_B(r) = \frac{1}{1 + \sum_{j \neq r} \exp(-(u_r - u_j))} = \frac{\exp(u_r)}{\sum_{j=1}^{m} \exp(u_j)} \tag{2.5}$$

There are several alternatives to the MNL, most of which are based on a different distribution assumption or the relaxation of the iid assumption. (ibid., ch. 8) Further on, the now motivated MNL is examined more thoroughly and the estimation process is explained in the succeeding chapter.

2.1.2 The Multinomial Logit Model as a Discrete Choice Model

Equation 2.5 corresponds directly to the generic form of the multinomial logit model with u_r as the predictor. (ibid., p. 227 f.) Other than within the classical linear regression, the predictor, hereafter written as η consists of two rather than one main components. Hereby, both the individuals as well as the categories have certain properties, possibly influencing the choice. The individuals' variables like age, sex or education are so-called global variables, while the category specific variables represents the quality of one component of the alternative at hand. Thinking about the common example of choice of modes of transport already used by (McFadden, 1973, p. 130 ff.), the category specific variables would be for example price or duration, while the individual specific are age or sex. In electoral choice, category specific variables are not as straight forward, as there is often no objective placement of parties and most scales on political topics are ambiguous. Thus, alternative specific variables are constructed. Given a position on a scale (like the left right spectrum in political science) each individual can subjectively place himself as well as all the participating parties on that scale. The absolute difference between the self position on the scale and the subjectively determined position of the party can now be interpreted as a property of the party for that individual, for which holds: the lower the absolute difference the better. Hereby, like in (Tutz, Pößnecker, and Uhlmann, 2015, p. 209) the alternatives are characterized by positions on policy dimensions, while the resulting difference can be seen as a natural scale. The absolute difference between the assessment and the self position is resulting in the variable $w_{ir} \in \{w_{i1}, \ldots, w_{ik}\}$ is therefore containing the attributes of category r for individual i.

The predictor resulting from individual and alternative specific properties for the category $r \in \{1, \ldots, k-1\}$ compared to the reference category k for individual $i \in \{1, \ldots n\}$ is thus written as:

$$\eta_{ir} = \beta_{r0} + x_i^T \beta_r + (w_{ir} - w_{ik})^T \alpha \qquad r = 1, \ldots, k-1 \tag{2.6}$$

β_{r0} and β_r are vectors containing the global parameters for the category r and α as the parameter determining the category specific influence, while $(w_{ir} - w_{ik})^T$ represent the difference to the reference category for individual i. Both parameter vectors are to be estimated using the underlying data.

Bringing this structuring of the predictor together with the model developed in equation 2.5 the form of log odds for choice of option r in comparison to the reference category k manifests itself as:

$$\log\left(\frac{P_B(r)}{P_B(k)}\right) = u_r - u_k = x^T \beta_r + (w_r - w_k)^T \alpha \tag{2.7}$$

with u_r and u_k being the utility towards the options r and k, while their difference can be written as the predictor specified above.

The multinomial logit model in its generic form in equation 2.5 is generally reliant on a side constraint to enable estimation. Here either choosing a reference category $k \in B$ resulting in $\beta_k = 0$ or using a symmetric constraint is necessary, while we rely on the former for further analysis.

2.1.3 Estimation in the Multinomial Logit Model

This chapter examines the estimation process for the parameter vectors β_r for each category $r \in \{1, \ldots, k-1\}$ and the vector α using maximum likelihood estimation.

The MNL can be written with regard to the multinominal distribution, which is an extension of the binomial distribution to multiple categories. (Fahrmeir, Kneib, and Lang, 2007, p. 237 ff.) therefore estimating parameters by determining the maximum of the likelihood (so-called maximum likelihood estimation) is possible. The most common approach, also used in this thesis, obtains the estimates by setting the first derivative of the log-likelihood to zero. (G. Young and Smith, 2005, ch. 8) The steps can be found in different literature, while we focus on the approach of (Tutz, 2011, ch. 8.6.1).

Let $\pi_1, \ldots, \pi_k$ be the probabilities related to the $1, \ldots, k$ options, while only $q = k-1$ probabilities are relevant as the kth probability can be calculated using the other ones. The y_i are multinomial distributed, with n_i repetitions with $k = q+1$ categories

$$(y_{i1}, ..., y_{iq}) \sim M(n_i, \pi_i) \tag{2.8}$$

following a multivariate exponential family.
Thus the density f of y_i can be written with regard to the multinomial distribution as:

$$\begin{aligned} f(y_i) &= \frac{n_i!}{y_{i1}!\cdots y_{iq}!(n_i - y_{i1} - \cdots - y_{iq})!}\pi_{i1}^{y_{i1}}\cdots\pi_{iq}^{y_{iq}}(1-\pi_{i1}-\cdots-\pi_{iq})^{(n_i-y_{i1}-\cdots-y_{iq})} \\ &= \exp(y_i^T\theta_i + n_i\log(1-\pi_{i1}-\cdots-\pi_{iq}) + \log(c_i)) \end{aligned} \tag{2.9}$$

This yields the likelihood

$$L(\beta) = \prod_{i=1}^{N} c_i\,\pi_{i1}^{y_{i1}}\cdots\pi_{iq}^{y_{iq}}(1-\pi_{i1}-\cdots-\pi_{iq})^{n_i-y_{i1}-\cdots-y_{iq}} \tag{2.10}$$

and thus the log-likelihood $l(\beta) = \sum_{i=1}^{N} l_i(\pi_i)$ with

$$l_i(\pi_i) = n_i\left\{\sum_{r=i}^{q} p_{ir}\log\Big(\frac{\pi_{ir}}{1-\pi_{i1}\cdots-\pi_{iq}}\Big) + \log(1-\pi_{i1}-\cdots-\pi_{iq})\right\} + \log(c_i) \tag{2.11}$$

The score function as the first derivative of the log-likelihood function, as well as the explicit Fischer information and covariance matrix of the maximum likelihood estimation in the MNL can be found in (Tutz, 2011, p. 219 f.). Further reading on maximum likelihood estimation and multinomial distributions is available for example in (G. Young and Smith, 2005, ch. 8).
The parametric model can therefore be estimated analytically and the uncertainty can be quantified for instance under a frequentist perspective.
There are three tests mainly used for the MNL: The Wald test, the Lagrange multiplier test and the Likelihood ratio test, see for example (Croissant, 2012, p. 24 ff.). Often, whether or not a specific estimate is different from zero is tested. Further reading on testing theory can be found for example in (G. Young and Smith, 2005, ch. 4) and applied for the MNL for example in (Tutz, 2011, p. 223 ff.)

Extension to semi-parametric models, hence splines and nonlinear fitting approaches are possible, see for example (Li, 2011) or (Kneib, Baumgartner, and Steiner, 2007), but in our case not practical as the number of degrees of freedom used by the model might exceed the desirable amount. Furthermore, the interpretation of splines is more complex. As we focus on the applicability of ontic imprecision in the realm of discrete choice models and the parametric model is by far more common and due to the arguments mentioned above we use the simple MNL as discrete choice model in this thesis.

2.1.4 Regularization and Variable Selection in the MNL

Variable selection is a crucial aspect of inference modeling if numerous potential variables are available, thus several approaches facing this problem have been discussed, see for example (Fahrmeir, Kneib, and Lang, 2007, ch. 3.6) or (Tutz, 2011, ch. 6). Here, in contrast to the mostly subjective and politological content related pre-selection in chapter 3.1.3, an ideally objective process is strived for. Irrelevant variables should be excluded from the model to ensure sparsity and model stability, as they widen prognoses intervals and the null hypothesis of an influence being zero tend to be mistakenly not rejected. (Fahrmeir, Kneib, and Lang, 2007, p. 156 ff.) There are several alternatives for model selection like the AIC, BIC or regularization. While AIC and BIC usually rely on stepwise functions, which leads to calculation problems with higher dimensional data and arbitrarily different results depending on the starting point, regularization has become more common among data scientists. (Mauerer et al., 2015, p. 26 f)

In this thesis we discuss variable selection by regularization, eventually following the approach of (Tutz, 2011, ch. 6 and 8).

Regularization is generally accomplished by maximizing a modified likelihood function consisting of the ordinary likelihood $l(\beta)$ and a penalization term $J(\beta)$ weighted multiplicative with the parameter λ. The goal can be to either improve model stability or variable selection by setting certain parameters to exactly zero. We will focus on the latter here. As in our case one variable consists of several parameters and the desired reduction is only achieved if every parameter of the variable is set to zero, we used grouped penalties like as proposed by (Tutz, Pößnecker, and Uhlmann, 2015, p. 209 f). While for ungrouped parameters the L1 norm is necessary to reduce parameter's values to exactly zero, (ibid., p. 210) chooses the L2 norm for the grouped parameters, which in this case shrinks the entire variable simultaneously to zero. This approach is also used by (Tutz, 2011, p. 235), in which it is explicitly argued that the L2 norm is sufficient as long as variables are grouped. Hereby, the L2 norm for the grouped parameters only, constitute of $\sum_{j=1}^{p} ||\beta_{.j}|| = \sum_{j=1}^{p} (\beta_{1j}^2 + \cdots + \beta_{k-1j}^2)^{1/2}$.

We follow (Tutz, Pößnecker, and Uhlmann, 2015, p. 2010) using the L2 norm with grouped classes for the global variables and the L1 norm for the ungrouped category specific variables in an MNL developed for discrete choice models about electoral research by (ibid.):

$$J(\beta) = \psi \sum_{j=1}^{p} \phi_j ||\beta_{.j}|| + (1-\psi) \sum_{l=1}^{L} \rho_l |\alpha_l| \tag{2.12}$$

This is called *Categorically Structured Lasso* (CATS), penalizing the vector $\beta_{.j}$ for each variable, as well as the category specific variables α_l, in its most general form. The penalization is internally weighted by ϕ_j for the global and ρ_l for the category specific variables. In order to obtain "fair" results using one λ value for both terms we choose $\phi_j = k-1$ and $\rho_l = 1$ following (Yuan and Lin, 2006, p. 51 ff.). The further tuning parameter $\Psi \in (0,1)$ to either weight regularization on the global or categorical estimates higher, is further on set to 0.5 as recommended by (Tutz, Pößnecker, and Uhlmann, 2015, p. 210). More about the idea behind equation 2.13 and further information about the CATS lasso can be found in (ibid.). As mentioned before the entire term $J(\beta)$ is weighted by the tuning parameter λ multiplicatively.

The success of the penalization essentially depends on the lambda value chosen, as it directly influences the process. While the penalized likelihood with a lambda value of zero is the ordinary likelihood, a high enough lambda value excludes all variables. As the "right" lambda value is not known, one usually conducts the regularization over a grid of lambda values, determining the most desirable by some other model criteria. Here either AIC, BIC or cross validation are common. (ibid., p. 211) While the BIC usually leads to very sparse models, cross validation usually favors a very small lambda. (ibid., p.219 f.) Thus using AIC is common practice and therefore also used in this analysis.

As argued in the introduction, we use regularization only for variable selection as a better alternative to step AIC or BIC algorithms, applying grouped classes to eliminate entire variables. After the unnecessary variables are determined, the model is refitted without regularization or these variables. The idea of refitting can be found in (Tutz, Pößnecker, and Uhlmann, 2015, p. 211), arguing that bias resulting from too high lambda values should be avoided.

2.2 Ontic and Epistemic Imprecision

This chapter provides the theoretical groundwork on how imprecision manifested in the form of a set can be interpreted regarding ontic or epistemic imprecision. In the succeeding chapter the application to discrete choice models is discussed.

There are two different ways to think about set valued data in the context of imprecision or incomplete information, discussed thoroughly for example in (Couso, Dubois, and Sánchez, 2014). On the one hand, one considers the set as an imprecise or fuzzy version of a true underlying value. On the other, the set or interval could be seen as an entity of itself.

The latter approach is called conjunctive or ontic imprecision, giving precise information of something imprecise or naturally set valued. Under ontic understanding, all viable information is contained in a set valued entity. For example the languages that an individual is capable of speaking, following (Couso and Dubois, 2014). Here, for example the set {German, English, French} contains all the information about one individual, the set cannot be reduced to one element and be more precise. The same applies if someone is truly undecided between options. The set containing all the options is the precise representation of the truth, as it is not possible to reduce the set to fewer elements by definition. This is an example for natural imprecision in its ontic form. To apply ontic imprecision in our case, there is a new state space needed, consisting of all possible combinations of the underlying units. This power set of the original specifications collects the new possible outcomes. As we have nominal, not ordinal data and a finite original space, we obtain a *finite random set* (Plass et al., 2015, p. 258) as the mapping from the original space

$$Y : \Omega \rightarrow P(S) \tag{2.13}$$

for any $A \subseteq S$ with S being the state space, P the power set and $(\Omega, \mathfrak{A})$ the underlying measurable space. Hence, all combinations are considered as possible outcomes, while for the inverse function holds $Y^{-1}(A) = \{\omega \in \Omega : Y(\omega) = A\} \in \mathfrak{A}$. (ibid., p. 258) Therefore any element within the combinations can be traced back to one element in the original set.

Thus, we have a measurable mapping on the power set, which allows us to conduct analysis, equivalently to the original state space consisting in our case of the singular parties. The new set is again of nominal scale and finite, while containing all set valued combinations descending from the original parties, providing the same basic mathematical properties like the original set. The resulting set valued elements of the power set can therefore be examined as categories of their own within for example choice modeling procedures. Hereby, the elements of the $P(S)$ constitute the subgroups of interest, while including both the single parties as sets, as well as all possible party combinations.

The other approach, called epistemic or disjunctive, yields a set in which the precise information is contained in. This means that one precise value lies within the imprecision of set valued information. If we change the question in the example above to: what language is most frequently used by the individual, the set {German, English, French} contains precisely one value which is accurate. If only this set is given as data, the goal has to be to reduce the imprecision which comes with the elements contained in the set. Usually a thorough understanding of the underlying structure has to be searched for, in order to retrieve or estimate this precise value. (Plass, 2018, p. 5 ff.) The epistemic approach again relies on the power set of the set of the original parties S: $\Omega \to P(S)$, only in this case the resulting combinations contain one true value from the original set. In mathematical terms the precise value can be understood as a function $Y_{precise} : \Omega \to S$. Equation 2.14 denotes the notation of the precise value lying in that set, resulting from this idea. The precise value $Y_{precise}(\omega)$ with ω being one element of the entire set Ω lies in the set valued information $Y_{epistemic}(\omega)$ which contains ω.

$$\{Y_{precise}(\omega) \in Y_{epistemic}(\omega), \forall \omega \in \Omega\} \tag{2.14}$$

This value contained in the set valued information hypothetically provided by the poll has to be estimated. For example, upper and lower bounds as one epistemic approach for predictions are discussed in (Plass et al., 2015, p. 260 f.)

More detailed information about epistemic and ontic understanding of set valued data and embedding in random sets and set theory can be found for example in (Couso, Dubois, and Sánchez, 2014), (Couso and Dubois, 2014) or (Nguyen, 2005).

2.3 Discrete Choice Models under Ontic Imprecision

Using set valued data with pre-election polls, regardless of being interpreted under ontic or epistemic imprecision, distinguish themselves from the conventional approach. For the conventional approach, used for example by (Tutz, Pößnecker, and Uhlmann, 2015) or (Mauerer et al., 2015), the individual is forced to state a precise singular party or is not considered in the analysis. Therefore, truly undecided individuals can not be included in the analysis, as well as no uncertainty related to the choice of the stated party of any participant is represented at all. The assumption that every individual is identically convinced about their choice seems very strong and would be weakened severely if set valued answers are introduced. The undecided individual is therefore either excluded from the analysis or chooses one precise, but inaccurate value. Especially with the background of an increasing number of undecided individuals thorough examination of this phenomena is of interest. Hereby, the set valued answer is the precise representation of the reality at that given point in time, regardless of how imprecise it is concerning the voting outcome. Furthermore, the degree of certainty is represented with the number of elements in the set, in contrast to the conventional approach, for which any precise statement is regarded the same. In this thesis, we thus emphasize the importance of the set valued approach.

For the set valued approach, both the epistemic and the ontic understanding of imprecision are applicable, even though specializing on different questions. For one, the disjunctive or epistemic approach seems plausible, as each individual is forced to cast one final, not set valued, vote on election day. Consideration of the final vote and therefore reduction of the set valued answers of undecided individuals is necessary for election outcome predictions. The underlying imprecision is obvious, as the real information (the party which the individual will vote for) is assumed to lie precisely in the set of the undecided individual. The process of how this vote comes about is not clear, and might occur due to a coin flip in the voting cabin or some other arbitrary procedure. It does not even have to be known by the individual itself yet, but for the epistemic view it is enough that this single value exists in the set of all possibilities. For discrete choice models, as introduced above, epistemic imprecision is not straight forward applicable, as the choice is mandatory for the estimation.
On the other hand, the ontic understanding of the set valued responses obtained lays emphasis on a different topic in political science, namely the one of voting behavior. Hereby, the undecided individuals pondering between specific parties are explicitly of interest. For both the affected parties and political scientists these individuals constitute further groups in the process of the voting decision process. The set of parties precisely determines their position. Thus, analyzing behavior of these groups using choice modeling seems evident. Considering these groups as entities of themselves provides a different angle on the decision processes.
This resembles the argument of (Oscarsson and Rosema, 2019) mentioned in the introduction, that the natural decision process includes a stage in which individuals are pondering between options (ibid., p. 257 ff.) and that this stage provides valuable information that should not simply be thrown away. Unlike their research, we further the approach by explicitly analyzing choice of the individual members in a group undecided between specific parties.
Hereby, we model the probability of individuals with certain features being in a specific group, decided or undecided. Thus, the discrete choice model does not focus on the election outcome, but on the underlying question of the real political position of the polls participants, contributing to an important topic in electoral research.

The state space of discrete choice models under ontic imprecision consists of the power set of the single parties, used as state space by conventional models, like as explained in the chapter above. The original nominal and finite alternatives given by the parties, in our case of the six main parties of Germany {AfD, FDP, CDU/CSU, SPD, Green, Left} [1] is therefore extended to the set of all the possible combinations of these six parties leading to 63 alternatives excluding the empty set. These are once again nominal, and finite options enabling straight forward discrete choice models. All common statistical methods applicable under the conventional approach can be used for the new state space.

[1] Alternative für Deutschland (AfD), Freie Demokratische Partei (FDP), Christlich Demokratische Union (CDU), Christlich Soziale Union (CSU), Sozialdemokratische Partei (SPD), Bündnis 90/Die Grünen (Green), Die Linke (Left)

3 Discrete Choice Models for the 2017 German Federal Election under Ontic Imprecision

In this chapter the methods developed in chapter 2 will be set to practice using the data from the *German Longitudinal Election Study* for the German federal election of 2017. First, the data's background, the pre-selection of variables, descriptive connections and the construction of the variable "ontic" will be discussed. Second, the models with and without ontic imprecision will be specified and third, results will be interpreted and models compared.

3.1 The Data from the German Longitudinal Election Study

In this chapter the data's background, as well as the essential, artificial construction of the ontic variable later on written as "ontic" is discussed. Furthermore, the qualitative pre-variable selection, to ensure comparability to recent research with the conventional approach is described and descriptive connections as well as item nonresponse is examined.

3.1.1 Background and Emergence of the Data

The *German Longitudinal Election Study* (GLES), founded in 2009, is a project in cooperation with several research facilities to provide continually high quality data for voting and election research in Germany, financed by the *German Research Foundation*. (GLES, 2019b) It is the most extensive German election survey and one of the biggest worldwide. (ibid.) The study was founded as a cross sectional survey for the federal elections of 2009, 2013 and 2017 with longitudinal character and will in future be extended and carried out in cooperation with DGfW. [2] (GLES, 2019a)
For each federal election a pre- and post-election survey is conducted with a sample size of about 2100, referring to all citizens over 15 years old who are eligible to vote. (GLES, 2019c) The proportion of participants from East to West Germany is $1/3$ to $2/3$ respectively, hence over representing East Germany, in order to enable in depth analysis of the important issue of geographical difference. Other than this, representativity is strived for within the sampling design. (ibid.) About one hour long Computer Assisted Personal Interviews (CAPI) are used to collect the data and in the case of the pre-election survey the interviews are conducted in the timespan beginning two months until few days before the election. In the timespan of 48 days, about 42 participants are interviewed every day.(ibid.) For our analysis we assume respondents to be homogeneous over the timespan and we do not adjust for the disproportional amount of East German citizens, following the referenced work using this source of data.
As we intend to apply ontic imprecision in our analysis, only the pre-election survey is used. Further, we only focus on the most recent election of 2017 in a cross sectional, non longitudinal manner for which 2179 observations are available with 2117 eligible to vote. The survey holds a huge variety of

[2] Deutsche Gesellschaft fÃijr Wahlforschung: German Society for Electoral Research

questions, selected and constructed by quantitative and qualitative research, resulting in 119 different questions in the pre-election survey alone. (GLES, 2019c)
Additionally to the usual question about the single preferred party, the participants are explicitly asked about their certainty attached to their intent as well as their assessment of the main participating parties. This enables us to construct a set valued variable "ontic" in chapter 3.1.2 and thus makes analysis under ontic imprecision possible with this dataset.
The six main parties in Germany above the necessary 5% threshold to enter the parliament are {AfD, FDP, CDU/CSU, SPD, Green, Left}, which are considered in the analysis, while the other parties are cumulated in an "other" category and are effectively not considered for the choice modeling. The parties CSU and CDU, competing in different states but operating as one party on federal level are combined.
In the German federal election system, there are two votes of which the first is cast for a politician (mostly member of a party), representing the election district while the second is cast for a party directly. Due to tactical voting for the first vote (Herrmann, 2015, ch. 3-5) and candidate dependency, we primarily focus on the second vote, following most research like (Mauerer et al., 2015) or (Tutz, Pößnecker, and Uhlmann, 2015). Further details on the German electoral system can be found under (Bundeswahlleiter, 2019).

3.1.2 Construction of the Variable "Ontic"

As mentioned in the introduction, this thesis contributes to the solution of a chicken-egg-dilemma (Fink, 2018, p. 7), by proving the applicability of discrete choice models under ontic imprecision with an artificial constructed variable "ontic". The construction, implemented in this chapter is therefore necessary. Further, we want to gain more detailed and accurate results about individuals' behavior for the German federal election of 2017, by using the information provided by the GLES in an ontic manner. The arbitrary component in the construction process is outweighed by the benefits of representation of uncertain individuals.

As mentioned in chapter 3.1.1, sufficient information is provided by the survey for a credible construction process. Hereby, the preferred party if stated, the assessment of the parties on a scale of +5 to -5 and the certainty about the preferred party in the four categories very certain, fairly certain, not certain and not certain at all, are represented in the data. If an individual does not provide information for both the intended party and consequently neither about the associated uncertainty, he is categorized as uncertain. This leads to the graph plot in figure 1.
The category of participants being very certain is the biggest one, but the not certain or not certain at all category accounts for about 38 % of the observations. As mentioned in the introduction, the variable "ontic" is introduced to adjust for the uncertain individuals, which according to plot 1 seems highly recommended, as exclusion of all uncertain individual distorts the overall picture.
The construction relies primarily on the uncertainty of the individual, while each individual is assigned

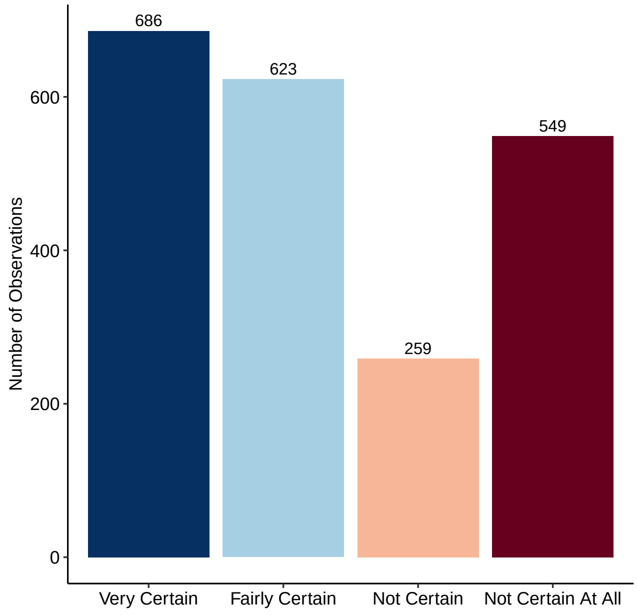

Figure 1: Frequencies of the degree of certainty attached to the intent within the data

a set containing at least one party. If the individual is not very certain, the set can be amended, while using the assessment of the parties. For each very certain individual, the stated party is the only one contained in the set, not differing from the conventional analysis. For the other individuals it is is distinguished between fairly certain and not certain or not certain at all. First, the participants who are fairly certain can be assigned more parties to their ontic set, if he or she assigns his maximal affinity to more than one party. Second, the set for the participants considering themselves not certain or not certain at all, any party will be amended that is assigned the highest or second highest rating by the individual on the scale between +5 and -5. At this point it can be argued for the approach of selecting all parties assessed positively. But as individuals tend to perceive scales very differently the individual specific highest and second highest rating seem more suited. This furthermore avoids dropping all individuals without any positive party assessment, which is not desirable.
Individuals confident to vote for none of the six big parties are excluded in the previous steps. We assume that at least most of the individuals who are uncertain and assess the big parties positively, will vote for one of them and not for one of the small parties.

For the MNL like for all other regression models, sufficient observations for each category is necessary to ensure stable estimation. Perfect separation has to be avoided, in the sense that estimation is not possible if there are no observations in one group for a certain variable. Thus we only permit categories with at least 40 observations leading to overall 13 categories of the possible 63 instead of the original six. The other categories unfortunately have to be left out leading to loss of observations. After the preprocessing we still have 1649 observations of the original 2117. Every observation for which the individual is not very certain and does not assess more than one party has to be excluded from the analysis. Figure 2 shows frequencies of the considered categories from left to right of the political

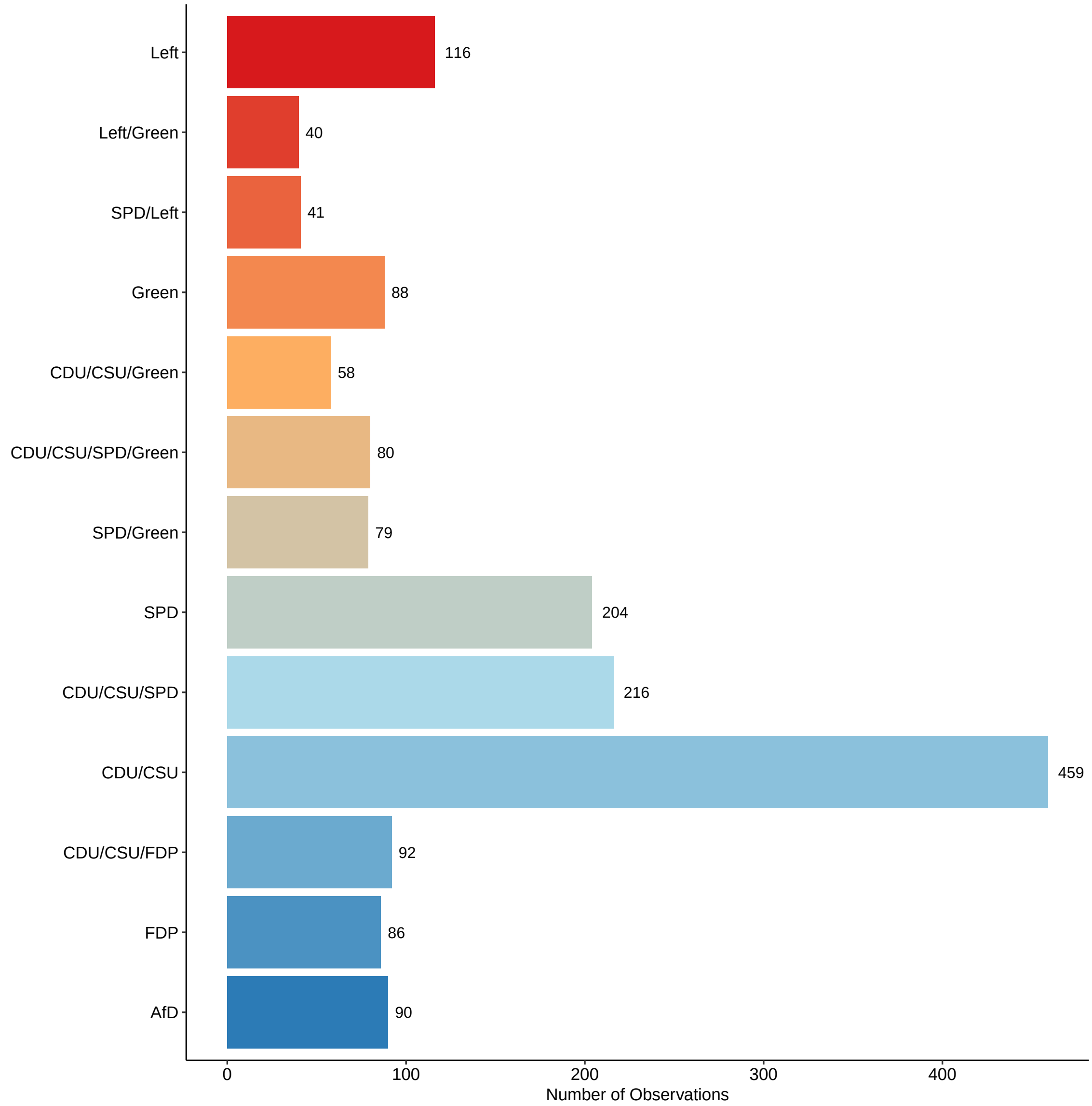

Figure 2: Number of observations for each constructed ontic category within the data

spectrum.
For further analysis all observations excluded for the constructed variable "ontic" are as well excluded from the conventional one, to ensure comparability. The connection between the variable "ontic" and the conventional single party variable and the resulting frequencies for the conventional approach are displayed in table 1 in chapter 3.1.4.

3.1.3 Pre-Selection and Variable Description

This chapter provides qualitative pre-variable selection for the succeeding models out of the 119 questions provided by the survey.
For reasons of clarity, avoidance of item nonresponse and model stability not all variables of the survey can be considered. To ensure comparability to other research without the novel ontic approach, we orientate us towards the recent publication on the German federal election using discrete choice models from (Mauerer et al., 2015), (Tutz, Pößnecker, and Uhlmann, 2015) and (Tutz, 2011). The aim of

the analysis is mostly to prove the concept, therefore direct comparison with recent research seems advisable.
Thus, in the first step 10 variables are selected, which according to (Mauerer et al., 2015, p. 31) are empirically and qualitatively justified. This results in the individual specific variables: age (numeric), sex (1:male, 0:female), most important source of information (1: internet, 0:other), member of a union (1:yes, 0:no), politically interested (1:yes, 0:no), resident from East or West Germany (1:West, 0:East), working class/lower class (1:yes, 0:no), confession christianity (1:yes, 0:no), satisfaction with democracy (1:yes, 0:no) and high school degree (1:yes, 0:no). For the year of 2017 the question of whether or not there should be a fixed maximum limit for refugees is amended, due to the emerging political importance of this question, at the time. (Haller, 2017, p. 22)
The variables concerning unemployment and union membership have to be left out, as the very low number of unemployed individuals and individuals in a union leads to perfect separation and thus estimation problems. As an alternative to unemployment the subjective social layer of the individual is considered.
For the category specific variables there are four options at hand, provided by the survey. Each relying on the subjective placement of oneself and all the parties considered on a certain scale. In the manner discussed in chapter 2.1.2 the absolute distance between those two points is used as a category specific variable. All placements of the individuals are subjective leading to a so called natural scale, for which only the personal distance is relevant while the objective or self position of the party is negligible. The possible variables provided by the survey are: The left/right position scale, the position toward climate change meaning how much the economy is allowed to suffer in order to reduce emissions, to what extent foreigners are allowed to move into Germany, and to what extend taxed should be increased to secure welfare and social projects. In order not to drop out of the analysis in an available case procedure, the individual has to place both himself as well as all parties on that scale. As this naturally, but regrettably occurs quite frequently depending on the question at hand, using all four variables would lead to vast reduction of sample size. On the other hand, do global variables usually have a lot of predictive power, and thus improve the analysis. As a compromise only the variable with by far the least drop outs is chosen, which is the left-right scale. For the set valued party combinations in the ontic analysis the mean of the differences towards the parties is used.
For the purpose of model stability and to improve the clarity concerning interpretation, most variables are regarded binary following the example of (Mauerer et al., 2015) and (Tutz, Pößnecker, and Uhlmann, 2015). This leads to sparse degrees of freedom used and enables straight forward interpretation. One example is the subjective socioeconomic status, which is originally divided in five categories: lower class, working class, lower middle, upper middle and upper class. This is modified into a binary variable whether or not the individual sees himself in the lower or working class. The age is a metric variable while categorical variables are generally avoided following the example of (Mauerer et al., 2015) and (Tutz, Pößnecker, and Uhlmann, 2015).
Figure 3 illustrates the nonresponse for every one of the randomly ordered 2117 individuals listed on the y-axis divided by each variable considered on the x-axis. Hereby, patterns of nonresponse can

be examined and the overall nonresponse situation is outlined. One can observe, that if one item is missing for a certain individual, another one is frequently missing as well.

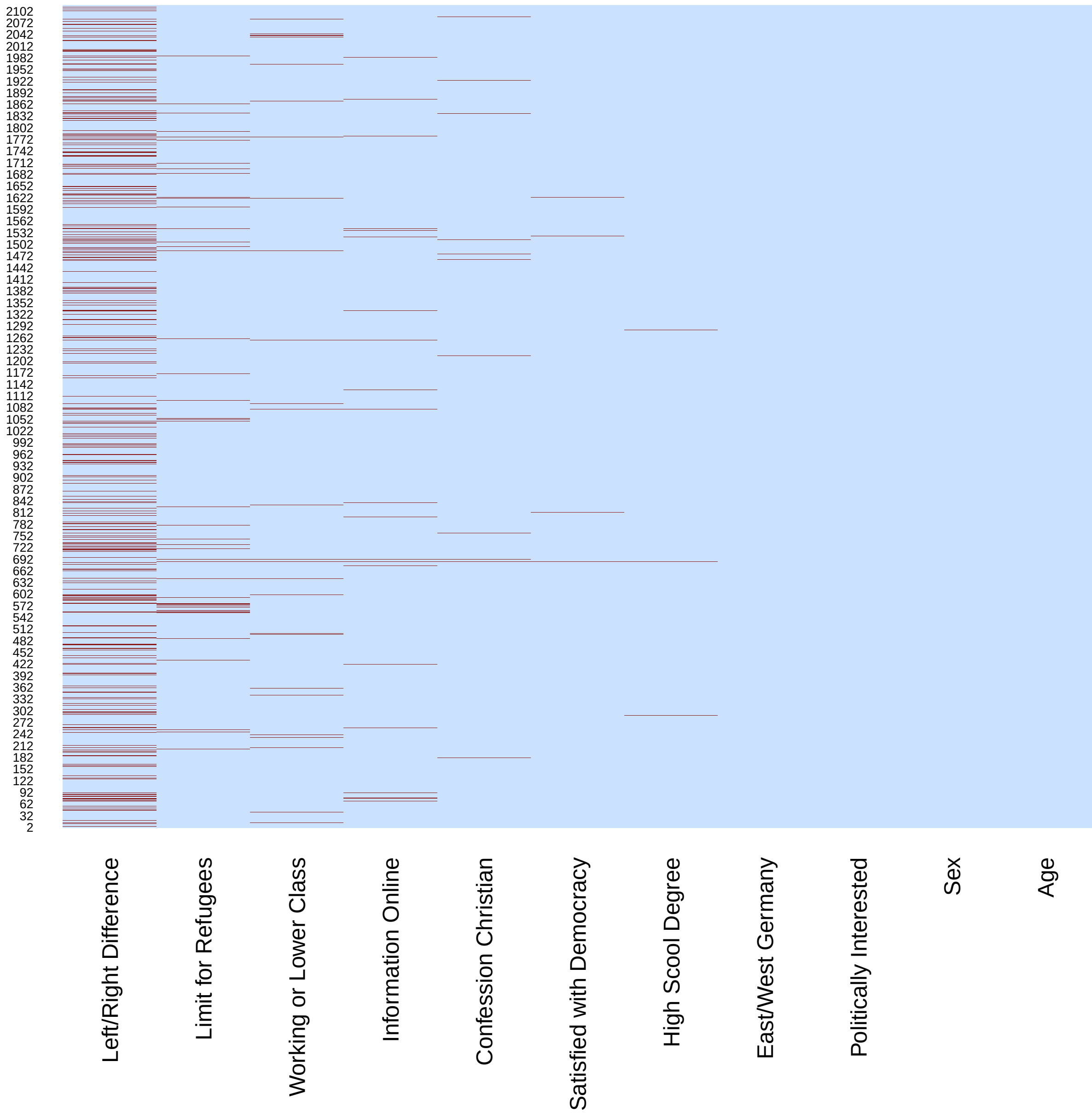

Figure 3: Illustration of the item nonresponse concerning the chosen variables in the data

Nonresponse is a considerable problem in any analysis using data from voluntary surveys. If the process behind the nonresponse is not random, bias is introduced. (Rubin, 2004) In this thesis we go for the most common way to deal with nonresponse by using the available cases after variable selection, fully aware of the drawbacks from this choice. But since all the reference work for this thesis also opts for available case analysis, we accept the disadvantages. For further works multiple imputation approaches, similar to what can be found in (ibid.) are possible.

3.1.4 Descriptive Statistics

In this chapter connections between the endogenous variables as well as properties of the exogenous variables, chosen in the previous chapter are examined.
As there are mostly binary variables, the problem of multicollinearity is not as delicate. The highest correlation between two independent variables is 35%, thus below the usual thresholds considered problematic.
Now we focus on the relation between the two endogenous variables for the models. Table 1 visualizes the frequencies of the dependent variables of both models. One can observe which observations of the "ontic" categories descent from which in the conventional model. Hereby, the "other" category represents the individuals that are undecided or did not choose one of the six main parties. As the

	AfD	FDP	CDU/ CSU/FDP	CDU/CSU	CDU/ CSU/SPD	CDU/ CSU/SPD/ Green	CDU/ CSU/ Green	SPD	SPD/ Green	SPD/Left	Green	Left/ Green	Left
AfD	85	0	0	0	0	0	0	0	0	0	0	0	0
FDP	0	85	35	0	0	0	0	0	0	0	0	0	0
CDU/CSU	0	0	16	452	55	14	15	0	0	0	0	0	0
SPD	0	0	0	0	62	11	0	199	20	13	0	0	0
Green	0	0	0	0	0	14	22	0	27	0	81	9	0
Left	0	0	0	0	0	0	0	0	0	14	0	11	113
Other	5	1	41	7	99	41	21	5	32	14	7	20	3

Table 1: Display of the connection between the conventional and ontic numbers of observations for each category

amount of certain individuals is rather high within the population, most individuals to list one specific party are in the set consisting of only this party for the constructed variable "ontic". The highest ratio of people not certain to vote for their preferred party choose the Green party, in contrast to individuals choosing the AfD of whom evidently everyone is very certain in this population. The "other" category, which is neglected for conventional analysis is assigned to different categories under ontic imprecision, with the main one being the set containing the most common parties CDU/CSU and SPD.

Further on, the descriptive connection between the binary exogenous and the endogenous variable for both models are displayed. Table 2 is a cross table between the frequencies of ones (the mean) for all binary variables and the ontic categories while table 3 shows them for the conventional approach. This provides the descriptive connection between the variables selected, except for the metric variable age and the global variable left/right scale. The results give a first impression of what to expect in the later analysis. Striking for both the ontic and the conventional table, is the high number of supporters for a fixed limit for refugees within the AfD supporters and their low level of satisfaction with democracy. People who are categorized to a set containing the green party or only the green party tend to be very satisfied with democracy.
There are also interesting distinctions between the ontic and conventional approach, with the ontic providing more detailed information. One example would be the proportions of christian individuals within the Left or Left/Green category. Hereby, the proportion of Left/Green christian voters is a lot

	Politically Interested	West/East Germany	Christian	Satisfied with Democracy	High School Degree	Limit for Refugees
AfD	0,389	0,422	0,378	0,111	0,233	0,871
FDP	0,523	0,686	0,547	0,721	0,512	0,607
CDU/CSU/FDP	0,413	0,783	0,648	0,565	0,467	0,703
CDU/CSU	0,375	0,734	0,706	0,693	0,373	0,661
CDU/CSU/SPD	0,296	0,639	0,623	0,574	0,37	0,696
CDU/CSU/Green	0,31	0,759	0,69	0,586	0,431	0,464
CDU/CSU/SPD/ Green	0,362	0,875	0,762	0,65	0,5	0,57
SPD	0,471	0,789	0,618	0,552	0,345	0,57
SPD/Green	0,418	0,873	0,692	0,633	0,633	0,367
Green	0,455	0,864	0,648	0,682	0,705	0,294
SPD/Left	0,39	0,415	0,415	0,39	0,512	0,4
Left/Green	0,575	0,7	0,525	0,5	0,625	0,35
Left	0,466	0,397	0,278	0,319	0,457	0,491

Table 2: The mean of the binary variables split up between the ontic categories

	Politically Interested	West/East Germany	Christian	Satisfied with Democracy	High School Degree	Limit for Refugees
AfD	0,412	0,447	0,376	0,118	0,247	0,862
FDP	0,517	0,733	0,563	0,683	0,508	0,641
CDU/CSU	0,364	0,748	0,721	0,692	0,388	0,654
SPD	0,433	0,744	0,613	0,566	0,362	0,577
Green	0,477	0,843	0,634	0,686	0,66	0,329
Left	0,486	0,413	0,328	0,341	0,478	0,471

Table 3: The mean of the binary variables split up between the conventional categories

higher than the one of the Left party alone, therefore more information is provided using the ontic approach which enables this more detailed statement. Several other cases can be found comparing the tables, showing that the ontic view generally provides more information. Furthermore, does the set containing only one party in the variable "ontic" sometimes differ from their counterpart in the conventional analysis, indicating differences between the approaches.

3.2 The Model

Within this chapter the MNL as a discrete choice model as explained in chapter 2.1 is applied to the data of the GLES for the German federal election in 2017. There are two models fitted to illustrate the differences between the two approaches, one relying on the set valued party combination chosen in chapter 3.1.2 and the other relying on the conventional approach. To ensure comparability, the models will be constructed using the same bases of observations and variables, thus excluding all observations for the conventional approach which are not considered for the ontic one due to item nonresponse. The

regularization is conducted and the results of both models are interpreted and compared.
First, regularization is used to determine the variables further used. Second, both models are finally explicitly specified and third interpreted, compared and evaluated.
The regularization is fitted using the R package `MRSP` from (Tutz, Pößnecker, and Uhlmann, 2015) and the final model with the package `mlogit` from (Croissant, 2012)

3.2.1 Applied Regularization

The regularization is used, as explained in chapter 2.1.4, as a tool for variable selection. Therefore, we conduct the regularization in order to determine the variables to use in the succeeding chapters.
We apply a categorically structured lasso also called CATS lasso, with a symmetric constraint to avoid an arbitrary influence by choosing a reference category. We use regularization following equation 2.12 in chapter 2.1.4 applied to the entire group of global estimates from one variable. Thus, if the influence of the entire variables' estimates is penalized to exactly zero the variable will no longer be regarded for further analysis. Furthermore, a grid of 100 lambdas is used for the fitting process, while the ideal lambda is determined by the resulting models' AIC. This procedure is motivated throughout (Tutz, Pößnecker, and Uhlmann, 2015) and (Mauerer et al., 2015, p. 30).
Figure 4 provides an overview of the regularization process and therefore which variables are permitted in the model. On the x-axis the $\log(1-\lambda)$ transformation of the lambda values is shown and dependent on this value the regularized estimates are illustrated. The line 1.5 determines the lambda value chosen by AIC model selection criteria. The group effect can clearly be observed, as the influences for one variable of all parties converge to exactly 0 with the same lambda value. For $\lambda = 0$ the plot shows the estimates for unpenalized discrete choice estimation for the model.
The variables for sex, politicly interested and working/lower class, are therefore being removed from the model by setting their coefficients to exactly zero. Thus, these variables will no longer be considered in the following model. If the same CATS model is employed under the conventional approach, no further variables drop out of the analysis.
Consequently, for further analysis a model with one category specific and seven individual specific variables is fitted for both the conventional and ontic approach. No further penalization is used for these models, following the argument of (Tutz, Pößnecker, and Uhlmann, 2015, p. 211) that refitting is preferable to ensure end results which are not too highly penalized and thus biased, as penalization always introduces some sort of bias.

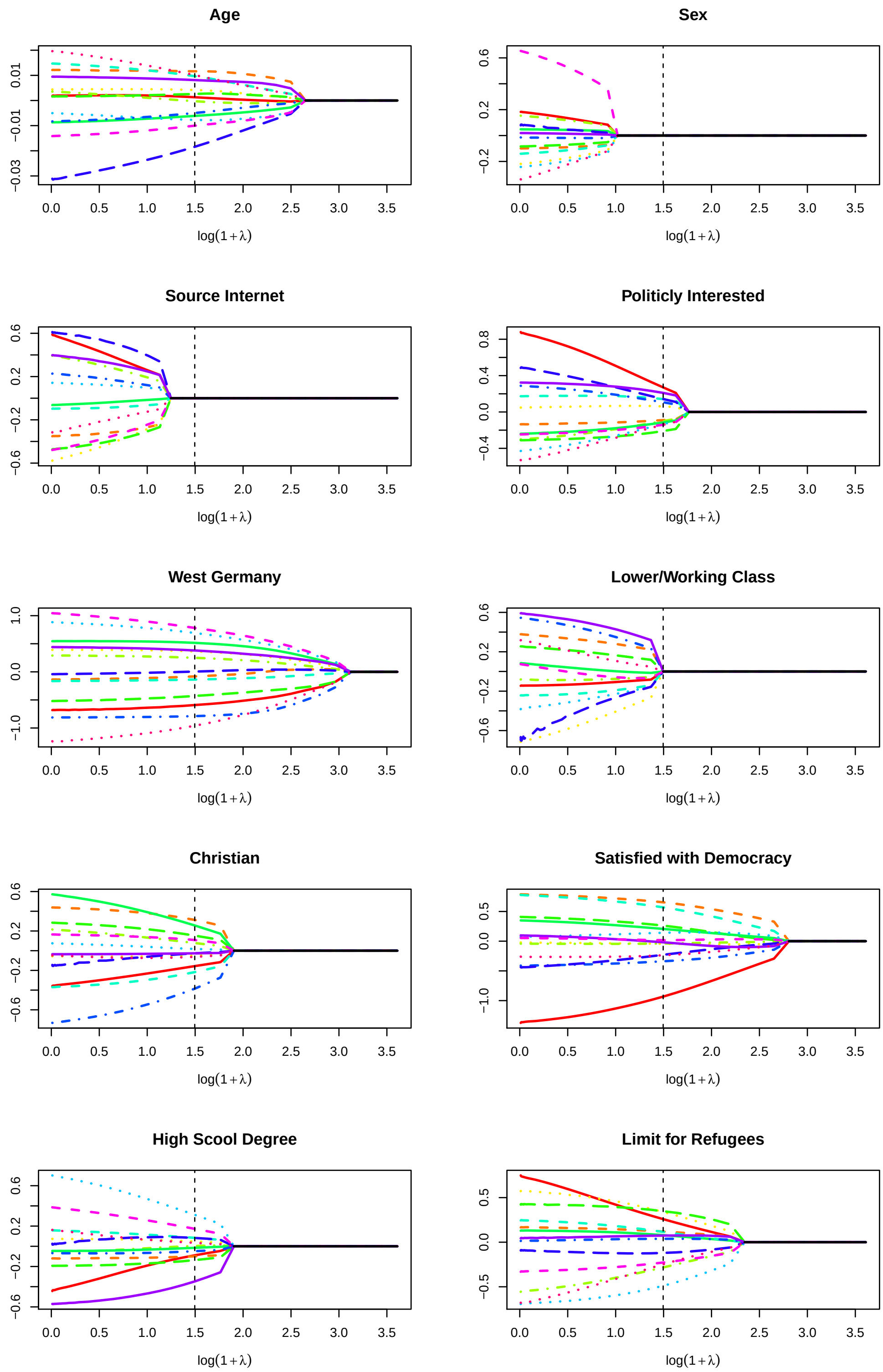

Figure 4: Visualization of the regularization applied to the binary variables chosen, dependent on the lambda value on the x-axis and the estimate on the y-axis

3.2.2 Model Specification

There are two final models specified in this chapter. To ensure comparability both models are based on the variable selection with one category specific and seven individual specific variables. For the same reason the observations are kept as close to each other as possible, leaving out the observations not in

the ontic approach for the conventional model.
The number of observations using available case analysis as explained in chapter 3.1.2 is 1390 for the ontic model while 1154 for the conventional. This difference results from the completely undecided voters that did not list a party but still are included in the ontic approach.
The predictor for both models is specified as:

$$\begin{aligned}\eta_r = \beta_0 + \alpha_{Left/Right} \cdot (w_{r,j} - w_{k,j}) + \\ \beta_1 \, x_{age} + \beta_2 \, I_{politically\ interested} + \beta_3 \, I_{west\ germany} + \\ \beta_4 \, I_{christian} + \beta_5 \, I_{satisfied\ w.\ democracy} + \beta_6 \, I_{high\ school\ degree} \\ \beta_7 \, I_{refugee\ limit}\end{aligned} \tag{3.1}$$

for the category r compared to reference category k, with $w_{r,j} - w_{k,j}$ denoting the difference to the reference category for the category specific influence for individual j. The indicator function for each binary variable is written as I, while the intercept is employed for model stability.
In order to solve the estimation problem some kind of restriction, usually in the form of a reference category has to be chosen. Results are always interpreted regarding this category. In our analysis the biggest party CDU/CSU, positioning itself in the middle of society [3], is chosen as reference category.

3.3 Results and Interpretation

In this chapter the models fitted according to the procedures explained in chapter 2 are diagnosed, interpreted and compared. The models are estimated using the R package `mlogit` developed by (Croissant, 2012). This chapter provides an outline of the most important results and compares the two approaches. The similarity is discussed, while some results are interpreted. As the underlying data and variables are the same with the exception of the truly undecided, the results of the two models can be directly compared. If the distinction between the undecided is not important, the results of the conventional approach should be similar to the results for the one party set contained in the ontic approach. On the other hand, if results differ, this emphasizes the virtue of the new approach, stressing the importance of including the uncertainty. As the uncertainty is accounted for, the new approach is more accurate than the conventional one, following both, a more nuanced and a more precise path. Therefore, even if a result is the same for both approaches, the new one can be seen as more reliable, as it does not depend on the more restrictive assumption concerning the undecided. For the analysis, after explaining how estimates are interpreted, we discuss the results of both models separately with primary focus on the new, ontic approach and compare the estimates contained in both models afterwards.

The results are interpreted as multiplicative odds, always referring to the reference category of the CDU/CSU under the ceteris paribus statement, thus only enabling relative, not absolute interpretation. For example, in table 4 for voters who say they are politically interested relative to those who are

[3] see: `https://www.cdu.de/system/tdf/media/dokumente/170703regierungsprogramm2017.pdf?file=1` (Last visited 16.10.2019)

not (first row, third column), the odds for voting for the AfD instead of voting for the CDU/CSU change multiplicative by the factor of $\exp(1,04) = 2.829$, given all other covariates stay the same. The category specific influence, only one estimate for each model, is interpreted analogously, by comparing one preference distance to another. Furthermore, it has to be noted that no statement about causality of the effect is made, but merely connections are displayed.
For each estimate the p-value from the test $H_0 : \beta_{i,j} = 0$ is provided, thus referring to an estimate significant different from zero. In tables 4 and 5 we only distinguish between $p < 0.1 : *; p < 0.05 : **; p < 0.01 : ***$ regarding the significance of influences. More on testing and how the statistic comes about can be found in (Croissant, 2012, p. 24 ff.).

First, looking at the ontic model, there are 104 individual specific and one category specific estimates. The results are illustrated in table 4, not showing the category specific estimate for the left/right difference with -0.756, which is significant with three stars. The model reaches a so called McFaddens R^2 of 0.20516, which constitutes a measure for the quality of fit for which each value between 0.2 and 0.4 is considered as an excellent fit. (McFadden, 1977, ch. 15) Furthermore, no estimation problems occur, as can be seen by sufficient values within the Hessian Matrix.
Roughly half the estimates are significantly different from zero at least for the level of 0.1. We primarily focus on those estimates to ensure at least some kind of reliability within the result. Regarding significance in such models with many variables, one has to be aware of the *multiple testing problem* as discussed by (Benjamini and Hochberg, 1995), according to which several estimates are possibly classified as significant even though they are not.

	Intercept	Age	Politically Interested	West/East Germany	Christian	Satisfied with Democracy	High School Degree	Limit for Refugees
AfD	1,018	-0,013	1,04 ***	-0,521	-0,798 **	-2,151 ***	-0,157	0,432
FDP	-1,724 ***	0,002	0,325	0,006	-0,821 ***	0,063	0,398	0,086
CDU/CSU/FDP	-1,271 **	-0,007	0,236	0,571 *	-0,463 *	-0,751 ***	0,313	0,39
CDU/CSU/SPD	0,474	-0,011 **	-0,185	-0,314	-0,162	-0,398 *	-0,071	0,255
CDU/CSU/SPD/Green	-0,838	-0,022 ***	-0,115	0,71 *	0,15	-0,427	0,135	-0,047
CDU/CSU/Green	-0,241	-0,015	-0,234	0,267	-0,263	-0,787 **	0,127	-0,734 **
SPD	-0,18	-0,007	0,407 *	0,563 **	-0,498 **	-0,716 ***	-0,495 **	-0,152
Green	-1,067	-0,02 **	-0,249	1,101 ***	-0,362	-0,685 **	0,924 ***	-0,897 ***
SPD/Green	-0,752	-0,025 ***	-0,218	1,196 ***	-0,191	-0,802 ***	0,54 *	-0,474
SPD/Left	-0,486	0,007	-0,371	-1,076 **	-0,533	-1,042 ***	0,286	-0,869 **
Left/Green	1,008	-0,051 ***	0,656	0,144	-0,555	-1,204 ***	0,322	-0,317
Left	1,757 ***	-0,024 ***	0,42	-0,679 **	-1,185 ***	-1,225 ***	-0,009	-0,145

Table 4: Results of the individual specific variables from the ontic model

Further on, from each column at least one estimate is interpreted and some remarks on possible implications on the political landscape are made.
An increase of age (second column) by one year decreases the chance to choose most parties compared to the CDU/CSU, hence the negative sign within all significant estimates. Especially, the chance to

choose the category Left/Green decreases with rising age and this, nearly twice as fast as with the categories containing the single parties Left or Green. Findings like these are only possible due to the ontic approach and would be dismissed within the conventional approach, not distinguishing between those groups of interest. Hereby, the already known tendency that young people rather vote for the Left or Green party (Brenke and Kritikos, 2017, p. 596 ff.) can be extended to the consideration that younger people are even more frequently undecided between those two parties. This is by all means an interesting finding.

The highest influence of subjectively being politically interested (third column) compared to those choosing CDU/CSU has the estimator of the AfD nearly tripling the chance. The AfD is widely considered as a protest party, for which a considerable number of individuals vote only to punish the established parties. (Hambauer and Mays, 2017, p. 135 ff.) Furthermore, the topics of the AfD are considered to be narrow by several sources (ibid.) and major political aspects like the position towards the retirement payments were not even addressed for a long time. This makes this result relying on subjective self placement even more interesting.

Differences between former West and East Germany (forth column) are an ongoing topic in electoral research due to the different cultural background and the disadvantageous economic situation in East Germany. (Bock and Fiedler, 2013) Our results confirm our expectation, that rather radical parties profit from the situation, namely the Left and the AfD. The set SPD/Left has the most negative effect, once again a finding, only possible with the ontic approach, while the chance for choosing the Green party or a set containing the Green party rises drastically in the West.

Not being of Christian faith (fifth column) reduces the chance for almost any other category than the CDU/CSU severely. Hereby, the word "christian" in both the CDU and CSU names probably plays a role. Surprisingly, the AfD has a rather high negative estimate with -0.798, even though the party presents itself as very concerned about the christian heritage as can be seen in their party program on page 11, 35 and 47 [4].

In comparison to the party CDU/CSU leading the government for the past 12 years at the time of this poll, most other categories induce not as much satisfaction with democracy (sixth column). Especially the AfD, a protest party to some extent, as mentioned before, with -2.151 provides the highest estimate of the analysis, signaling strong dissatisfaction.

For higher education (seventh column) the Green party holds the highest positive effect, which is a phenomenon frequently discussed in political post election analysis like in (Brenke and Kritikos, 2017, p. 599).

The maximal limit of refugees (eighth column) was originally proposed by a politician of the CSU. The estimates for categories containing the Green party and the Green party itself evidently signal opposition to that idea. Especially, the set of CDU/CSU/Green is interesting, as the chance decreases severely even though the party proposing the concept are within this category. This distinction is only possible due to the ontic analysis, providing more precise and thorough information.

[4] https://www.afd.de/wp-content/uploads/sites/111/2017/06/2017-06-01_AfD-Bundestagswahlprogramm_Onlinefassung.pdf (Last visited 16.10.2019)

While most results seem to concede with the public political opinion, very interesting points can be found within this model. Especially the extension to choice sets provides valuable, precise and new information, improving the conventional approach with a more differentiated view.

The conventional model relies on 48 individual and one category specific three star significant estimate of -0.670. The results can be seen in table 5 showing the 48 individual specific estimates in the seven categories plus intercept. The conventional model even reaches an McFaddens R^2 of 0.27838 signaling an excellent fit and also no occurrence of estimation problems occur.
In this thesis the results of the conventional modeling approach merely serves as comparison to the new, ontic, set valued approach and are therefore not discussed more thoroughly themselves.

	Intercept	Age	Politically Interested	West/East Germany	Christian	Satisfied with Democracy	High School Degree	Limit for Refugees
AfD	1	-0,015	0,946 ***	-0,328	-1,026 ***	-2,033 ***	-0,237	0,319
FDP	-1,573 ***	-0,001	0,441 *	0,319	-0,888 ***	-0,167	0,402	0,258
SPD	0,247	-0,007	0,229	0,293	-0,55 ***	-0,553 ***	-0,382 *	-0,105
Green	-0,245	-0,024 ***	0,154	0,842 ***	-0,55 **	-0,479 *	0,567 **	-0,685 ***
Left	1,402 **	-0,015 *	0,511 *	-0,646 **	-1,088 ***	-1,189 ***	-0,028	-0,417

Table 5: Results of the individual specific variables from the conventional model

The results of the two approaches will now be directly compared in table 6, examining differences and similarities for the six main parties contained between both models.

Variable	Intercept		Age		Politically Interested		West/East Germany		Christian		Satisfied with Democracy		High School Degree		Limit for Refugees	
Model	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**	**Ontic**	**Old**
AfD	1,018	1	-0,013	-0,015	1,04 ***	0,946 ***	-0,521	-0,328	-0,798 **	-1,026 ***	-2,151 ***	-2,033 ***	-0,157	-0,237	0,432	0,319
FDP	-1,724 ***	-1,573 ***	0,002	-0,001	0,325	0,441 *	0,006	0,319	-0,821 ***	-0,888 ***	0,063	-0,167	0,398	0,402	0,086	0,258
SPD	-0,18	0,247	-0,007	-0,007	0,407 *	0,229	0,563 **	0,293	-0,498 **	-0,55 ***	-0,716 ***	-0,553 ***	-0,495 **	-0,382 *	-0,152	-0,105
Green	-1,067	-0,245	-0,02 **	-0,024 ***	-0,249	0,154	1,101 ***	0,842 ***	-0,362	-0,55 **	-0,685 **	-0,479 *	0,924 ***	0,567 **	-0,897 ***	-0,685 ***
Left	1,757 ***	1,402 **	-0,024 ***	-0,015 *	0,42	0,511 *	-0,679 **	-0,646 **	-1,185 ***	-1,088 ***	-1,225 ***	-1,189 ***	-0,009	-0,028	-0,145	-0,417

Table 6: Comparison of the results of the individual specific variables from ontic and conventional models for the main six parties for the second vote

Overall the results are similar in most cases, but differences can be found. The Green Party diverges the most, which can be lead back to the highest uncertainty within this group as was displayed in chapter 3.1.4 in figure 1. Hereby, total differences of 0.357 in the case of High School Degree or 0.212 with the maximal limit for refugees arise. But also differences for other parties can be found like for the SPD and the satisfaction with democracy.
Even though differences do occur, the estimates reciprocally lie in the 95% confidence intervals of the other approach. This can probably be reduced to the fairly high uncertainty attached to the multinomial estimation process.

Due to these facts and most estimates being similar, apparently no severe bias is introduced by neglecting the uncertainty associated with the preferred party. Never the less, more nuanced and precise information is obtained in several cases, providing new insights into electoral choice for the German federal election of 2017. As argued before, the ontic model does not rely on the severe assumption concerning the undecided and can therefore be seen as the more accurate approach. The more credible results come at the price of more and thus smaller category groups in regard to sample size and hence more numerical instability. Overall the virtues of the new approach are evident, while the loss of numerical stability has to be taken into account.

In this paragraph the first vote is briefly analyzed regarding the differences between the ontic and the conventional model. The results for the first vote should be analyzed cautiously, as the first vote is prone to strategic voting (Herrmann, 2015) leading to difficulties with aggregated numbers and is as well person as party dependent. Never the less the results are shown following the idea of the proof of concept in table 7.

Variable	Intercept		Age		Politically Interested		West/East Germany		Christian		Satisfied with Democracy		High School Degree		Limit for Refugees	
Model	Ontic	Old	Ontic	Old	Ontic	Old	Ontic	Old	Ontic	Old	Ontic	Old	Ontic	Old	Ontic	Old
AfD	1,307 *	1,1	-0,736 ***	-0,025 **	-0,037	0,817 **	-0,572 *	-0,411	0,644 ***	-1,075 ***	-0,883 **	-2,392 ***	-1,288 ***	-0,315	0,965 ***	0,733
FDP	-0,592	-1,506 **	-0,02 **	-0,011	-0,229	0,557 *	-0,323	-0,218	-0,288	-0,853 ***	-0,854 ***	-0,752 **	-0,342	0,842 ***	0,532	0,443
SPD	-0,157	-0,19	-0,056 ***	-0,005	0,518	0,077	-0,713 **	0,485 **	-0,402	-0,551 ***	-0,503	-0,648 ***	0,269	-0,04	0,365 *	-0,032
Green	-0,619	-1,113 *	-0,021 **	-0,018 **	0,489	-0,179	0,893 **	0,836 **	-0,241	-0,625 **	-0,505	-0,689 **	0,595 **	1,078 ***	0,758	-0,464 *
Left	1,579 ***	1,119 *	-0,03 ***	-0,021 **	-0,124	-0,019	-0,369	-0,659 **	0,129	-1,147 ***	-0,438 **	-1,313 ***	0,449	0,585 *	0,267	-0,01

Table 7: Comparison of the results of the individual specific variables from ontic and conventional models for the main six parties for the first vote

The models unsurprisingly differ more than the ones for the second vote, due to reasons mentioned above with several changes in sign like for being of Christian faith for the AfD. These severe differences, even though regarded with caution, stress once again how the new approach provides a distinguishable perspective.

4 Concluding Remarks

This thesis concerned itself with a method of choice modeling taking into account the undecided voters in multiparty systems. While the conventional approach, only providing the option of a precise answer falls short to deal with the rising number of undecided individuals, set valued response constitutes an opportunity to face this problem. The resulting set valued information can be analyzed under ontic or epistemic imprecision, while for an examination of groups undecided between specific parties prior to the election the ontic approach is advisable. We demonstrated that state of the art choice modeling is applicable under ontic imprecision, by extending the original state space to its power set. Following this idea, all statistical methods for the original, nominal state space can be transferred to the new power-set, providing methodological groundwork for further research on this topic.

To conduct the analysis for the German federal election of 2017, the set valued response was artificially constructed using the uncertainty and assessment of parties available for each individual. The results show similarity between most estimates of the conventional and ontic approach, but some differences can be found. Hereby, the ontic approach not reliant on the strong assumption concerning the undecided, provides more detailed information and directly addresses the groups undecided between specific parties. The approach is therefore advantageous in two ways: One, laying the methodological foundation for further research under ontic imprecision and two, providing more detailed, plausible choice modeling for the German federal election of 2017 by addressing the undecided.

Now that the methodological applicability is proven to be proficient, we strongly suggest to allow for multiple answers in pre-election polls, collecting valuable data and directly allowing for the ontic approach. One drawback of the approach is the smaller number of observations in each group if multiple categories are amended without enlargement of the overall sample size, increasing numerical instability within the new approach. On the other hand are undecided individuals not neglected and therefore more information is used. Nevertheless this consideration process has to be evaluated carefully.

Set valued information, gathered in pre-election polls opens the door to all kinds of interesting, new research. Epistemic approaches for bayesian electoral outcome predictions as well as further choice modeling and ontic imprecision are promising. As now methods for dealing with set valued response are available, this idea could impact all electoral research, leading to analysis not relying on a too strong assumption in times of increasing numbers of undecided voters.

References

Adams, J., S. Merrill III, and B. Grofman. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. Cambridge University Press, 2005.

Benjamini, Y. and Y. Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society: series B (Methodological)* 57.1 (1995), pp. 289–300.

Bock, K. and W. Fiedler. *Umbruch in Ostdeutschland: Politik, Utopie und Biographie im Übergang*. Springer-Verlag, 2013.

Brenke, K. and A. Kritikos. "Wählerstruktur im Wandel". In: *DIW-Wochenbericht* 84.29 (2017), pp. 595–606.

Bundeswahlleiter. *Das Wahlsystem*. (Last visited 14.10.2019). 2019. URL: `https://www.bundeswahlleiter.de/bundestagswahlen/2017/informationen-waehler/wahlsystem.html#92c6d030-288d-4dea-bd8d-3877740358f2`.

Couso, I. and D. Dubois. "Statistical Reasoning with Set-Valued Information: Ontic vs. Epistemic Views". In: *International Journal of Approximate Reasoning* 55.7 (2014), pp. 1502–1518.

Couso, I., D. Dubois, and L. Sánchez. "Random Sets and Random Fuzzy Sets as Ill-Perceived Random Variables". In: *Springer Briefs in Computational Intelligence*. Springer, 2014.

Croissant, Y. "Estimation of Multinomial Logit Models in R: The mlogit Packages". In: *R package version 0.2-2* (2012).

Eijk, C. van der, W. van der Brug, M. Kroh, and M. Franklin. "Rethinking the Dependent Variable in Voting Behavior: On the Measurement and Analysis of Electoral Utilities". In: *Electoral Studies* 25.3 (Sept. 2006), pp. 424–447.

Fahrmeir, L., T. Kneib, and S. Lang. *Regression*. Springer, 2007.

Fenwick, I., F. Wiseman, J. Becker, and J. Heiman. "Classifying Undecided Voters in Pre-Election Polls". In: *Public Opinion Quarterly* 46.3 (1982), pp. 383–391.

Fink, P. "Contributions to Reasoning on Imprecise Data". PhD thesis. LMU, 2018.

Fredén, A. and J. Sohlberg. "Insurance Strategic Considerations in Coalition-Oriented Systems: A Consideration Set Model Approach". In: *Electoral Studies* 57 (Feb. 2019), pp. 302–308.

GLES. *Das GLES-Design im Überblick*. (Last visited 02.10.2019). 2019. URL: `http://gles.eu/wordpress/design/`.

— *German Longitudinal Election Study*. (Last visited 01.10.2019). 2019. URL: `https://www.gesis.org/wahlen/gles/`.

— *Vor- und Nachwahl-Querschnitt*. (Last visited 13.10.2019). 2019. URL: `http://gles.eu/wordpress/design/querschnitt/`.

Haller, M. "Die Flüchtlingskrise in den Medien". In: *Frankfurt Am Main: Otto Brenner Stiftung* (2017).

Hambauer, V. and A. Mays. "Wer wählt die AfD? – Ein Vergleich der Sozialstruktur, politischen Einstellungen und Einstellungen zu Flüchtlingen zwischen AfD-WählerInnen und der WählerInnen der anderen Parteien”. In: *Zeitschrift für Vergleichende Politikwissenschaft* 12.1 (Dec. 2017), pp. 133–154.

Hasan, A., Z. Wang, and A. Mahani. “Fast Estimation of Multinomial Logit Models: R Package mnlogit”. In: *Journal of Statistical Software* 75.3 (2016).

Herrmann, M. *Strategisches Wählen in Deutschland*. Springer Fachmedien Wiesbaden, 2015.

Jennings, W. and C. Wlezien. “Election Polling Errors Across Time and Space”. In: *Nature Human Behaviour* 2.4 (Mar. 2018), pp. 276–283.

Keeter, S. “Are Public Opinion Polls Doomed?” In: *Nature Human Behaviour* 2.4 (Mar. 2018), pp. 246–247.

Kneib, T., B. Baumgartner, and W. Steiner. “Semiparametric Multinomial Logit Models for Analysing Consumer Choice Behaviour”. In: *AStA Advances in Statistical Analysis* 91.3 (2007), pp. 225–244.

Li, B. “The Multinomial Logit Model Revisited: A Semi-Parametric Approach in Discrete Choice Analysis”. In: *Transportation Research Part B: Methodological* 45.3 (2011), pp. 461–473.

Mauerer, I., W. Pößnecker, P. Thurner, and G. Tutz. “Modeling Electoral Choices in Multiparty Systems with High-Dimensional Data: A Regularized Selection of Parameters using the Lasso Approach”. In: *Journal of choice modelling* 16 (2015), pp. 23–42.

McFadden, D. “Conditional Logit Analysis of Qualitative Choice Behavior”. In: *Zarembka, P., Ed., Frontiers in Econometrics, Academic Press, New York* (1973), pp. 105–142.

— *Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments*. Cowles Foundation Discussion Papers 474. Cowles Foundation for Research in Economics, Yale University, 1977.

Nguyen, H. “Fuzzy and Random Sets”. In: *Fuzzy Sets and Systems* 156.3 (Dec. 2005), pp. 349–356.

Oscarsson, H. and M. Oskarson. “Sequential Vote Choice: Applying a Consideration Set Model of Heterogeneous Decision Processes”. In: *Electoral Studies* 57 (Feb. 2019), pp. 275–283.

Oscarsson, H. and M. Rosema. “Consideration Set Models of Electoral Choice: Theory, Method, and Application”. In: *Electoral Studies* 57 (Feb. 2019), pp. 256–262.

Plass, J. “Statistical Modeling of Categorical Data under Ontic and Epistemic Imprecision”. PhD thesis. LMU, 2018.

Plass, J., P. Fink, N. Schöning, and T. Augustin. “Statistical Modelling in Surveys without Neglecting "The Undecided": Multinomial Logistic Regression Models and Imprecise Classification Trees under Ontic Data Imprecision - extended version”. In: *Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications* (2015), pp. 257–266.

Rekker, R. and M. Rosema. “How (Often) do Voters Change their Consideration Sets?” In: *Electoral Studies* 57 (Feb. 2019), pp. 284–293.

Rubin, D. *Multiple Imputation for Nonresponse in Surveys*. Vol. 81. John Wiley & Sons, 2004.

Steinbrecher, M. and H. Schoen. “Channelling Attention and Choice? Examining Effects of Consideration Sets on Electoral Decision-Making”. In: *Electoral Studies* 57 (Feb. 2019), pp. 294–301.

Thurner, P. “The Empirical Application of the Spatial Theory of Voting in Multiparty Systems with Random Utility Models”. In: *Electoral Studies* 19.4 (2000), pp. 493–517.

Train, K. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2009.

Tutz, G. *Regression for Categorical Data*. Vol. 34. Cambridge University Press, 2011.

Tutz, G. and J. Gertheiss. “Regularized Regression for Categorical Data”. In: *Statistical Modelling* 16.3 (2016), pp. 161–200.

Tutz, G., W. Pößnecker, and L. Uhlmann. “Variable Selection in General Multinomial Logit Models”. In: *Computational Statistics & Data Analysis* 82 (2015), pp. 207–222.

Vassil, K., M. Solvak, and P. Ehin. “More Choice, Higher Turnout? The Impact of Consideration Set Size and Homogeneity on Political Participation”. In: *Journal of Elections, Public Opinion and Parties* 26.1 (Nov. 2015), pp. 78–95.

Young, G. and R. Smith. *Essentials of Statistical Inference*. Cambridge University Press, 2005.

Young, M. “The Problem with Categorical Thinking by Psychologists”. In: *Behavioural Processes* 123 (Feb. 2016), pp. 43–53.

Yuan, M. and Y. Lin. “Model Selection and Estimation in Regression with Grouped Variables”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68.1 (Feb. 2006), pp. 49–67.

Zeit. *Die Hälfte der Wähler hat sich noch nicht entschieden*. Die Zeit: Online Newspaper. (Last visited 14.10.2019). Aug. 2017. URL: https://www.zeit.de/politik/deutschland/2017-08/bundestagswahl-umfrage-waehler-unentschlossen.