The Journal of Choice Modelling 16 (2015) 23–42

Contents lists available at ScienceDirect

The Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Modeling electoral choices in multiparty systems with high-dimensional data: A regularized selection of parameters using the lasso approach

Ingrid Mauerer [a,*], Wolfgang Pößnecker [b], Paul W. Thurner [a], Gerhard Tutz [b]

[a] Department of Political Science, Ludwig Maximilian University of Munich, Oettingenstr. 67, D-80538 Munich, Germany
[b] Department of Statistics, Ludwig Maximilian University of Munich, Akademiestr. 1, D-80799 Munich, Germany

ARTICLE INFO

Article history:
Received 21 December 2014
Received in revised form
14 August 2015
Accepted 20 September 2015
Available online 6 October 2015

Keywords:
Parameter selection
Lasso
Multinomial logit model
Multiparty elections

ABSTRACT

The increased usage of discrete choice models in the analysis of multiparty elections faces one severe challenge: the proliferation of parameters, resulting in high-dimensional and difficult-to-interpret models. For example, the application of a multinomial logit model in a party system with J parties results in maximally $J-1$ parameters for chooser-specific attributes (e.g., sex and age). For the specification of alternative-specific attributes (usually: positions on issues and issue distances), maximally J parameters for each political issue can be estimated. Thus, a model of party choice with five parties based on three political issues and ten voter attributes already produces 59 possible coefficients. As soon as we allow for interaction effects to detect segment-specific reactions to issues, the situation is even aggravated. In order to systematically and efficiently identify relevant predictors in voting models, we derive and use Lasso-type regularized parameter selection techniques that take into account both individual- and alternative-specific variables. Most importantly, our new algorithm can handle for the first time the alternative-wise specification of the attributes of alternatives. Applying the specifically adjusted Lasso method to the 2009 German Parliamentary Election, we demonstrate that our approach massively reduces the models' complexity and simplifies their interpretation. Lasso-penalization clearly outperforms the simple ML estimator. The results are illustrated by innovative visualization methods, the so-called effect star plots.

1. Introduction

The usage of discrete choice models in the analysis of voters' choice behavior in multiparty systems has become common practice. Multinomial logit and probit models[1] are well-established tools in electoral research. Extensive analyses based on voter characteristics (age, sex, income etc.) on one hand and attributes of the parties or candidates on the other hand[2]

* Corresponding author.
E-mail addresses: Ingrid.Mauerer@gsi.uni-muenchen.de (I. Mauerer), wolfgang.poessnecker@stat.uni-muenchen.de (W. Pößnecker), paul.thurner@gsi.uni-muenchen.de (P.W. Thurner), gerhard.tutz@stat.uni-muenchen.de (G. Tutz).

[1] In the following we use the general term "multinomial logit model" to refer to multinomial logit models including both alternative- and individual-specific variables. In addition, we use the terms "alternative-specific", "choice-specific" and "party-specific" covariates interchangeably to refer to attributes of alternatives (i.e., parties/candidates). Analogously, "chooser-specific", "voter-specific" or "individual-specific" variables refer to voter characteristics.

[2] For example, the distance between voter's most preferred position on scales of controversial issues and the corresponding position of these

http://dx.doi.org/10.1016/j.jocm.2015.09.004
1755-5345/

contributed enormously to our understanding of the strategies of party competition (such as position-taking and issue saliency) and of voters' individual choice behavior (see Alvarez and Nagler, 1998; Adams et al., 2005). However, the well-known flexibility of discrete choice models comes along with a severe drawback: the proliferation of possible coefficients. This is the reason why we propose for the first time to consider sophisticated parameter selection techniques and to adjust them to this specific context of discrete choice modeling. To illustrate the problem more precisely, party choice in a party system with J parties results in maximally $J-1$ parameters in the case of chooser-specific attributes. For each alternative-specific attribute, maximally J parameters can be estimated. As a consequence, the amount of possible individual- and alternative-specific coefficients increases rapidly, resulting in highly complex and difficult-to-interpret models. Moreover, as soon as we allow for interaction effects, e.g., in order to test for segment-specific reactions to issue distances, the situation is even aggravated. For these reasons, the following research question is raised: how can we systematically and efficiently reduce this high-dimensional parameter space?

In this paper, we address the problem of parameter proliferation in discrete choice modeling. Our objective is to efficiently model (electoral) choices based on very high-dimensional data. We introduce and derive a Lasso-type regularization technique in the estimation of multinomial logit models (MNLs) which takes into account both individual- and alternative-specific variables. Most importantly, our new tailor made solution allows us to handle for the first time the alternative-wise specification of choice-specific covariates[3] (e.g., issue distances).[4]

The Lasso approach is a regularized parameter selection technique that penalizes the L_1-norm of the coefficients. This enforces parameter selection and reduction of the predictor space, and therefore guarantees – in contrast to classical subset selection approaches – continuous, stable and computationally efficient variable selection.

By applying the Lasso method to the 2009 German Parliamentary Election, we demonstrate that our proposed approach massively reduces the model's complexity and simplifies its interpretation by selecting the most important determinants of party choice. In addition, we show that Lasso-penalization improves the model's predictive performance. More specifically, we demonstrate that Lasso-penalization clearly outperforms the simple ML estimator, for both a main effects' and an interaction effects' model. The results are summarized by using the so-called effect star plots.

The contribution of this paper is threefold. First, it introduces and illustrates the benefits of Lasso-type regularization techniques in the study of choice behavior, and therefore offers a highly promising solution to the problem of high-dimensional predictor spaces in MNLs. Second, in contrast to previous work, our new Lasso method not only takes into account both individual- and alternative-specific variables, but also is able to handle the alternative-wise specification of choice-specific variables. Even though the benefits of our Lasso approach are illustrated by using electoral choices, it can be applied to any choice situation based on this type of model class facing the problem of parameter inflation.[5] Third, we show that using a symmetric identifiability constraint prevents that the arbitrary choice of a reference category biases the Lasso's selection of the most promising model.

The remainder of the paper is structured as follows: to introduce the problem of high-dimensional data in modeling electoral choices in multiparty systems, Section 2 provides a short formal outline of the theoretical model. Then, we briefly examine classical variable selection procedures and their limitations. In particular, we demonstrate why and how the usage of regularization methods in the study of multiparty elections enables us to efficiently identify important predictors, and therefore to improve electoral choice models, and to facilitate their interpretation. In Section 4 we illustrate the usefulness of the proposed approach by providing a regularized analysis of party choice in the 2009 German Parliamentary Election. Finally, Section 5 summarizes the major advantages of our approach in the analysis of electoral choice behavior.

2. Model formulation

MNLs are mainly used in the study of individual voting behavior to translate the Spatial Theory of Voting into a statistical model.[6] The Spatial Theory of Voting assumes that voter i chooses the party/candidate j that offers policy positions on K policy dimensions, denoted by p_{ijk}, that are closest to the voter's most preferred policy position on each k dimension, denoted by x_{ik}. According to the principle of utility maximization, voter i compares the parties' policy proposals p_{ijk}[7] and

(*footnote continued*)
alternatives.

[3] In contrast to Tutz et al. (2015), who propose the so-called Categorically Structured Lasso (CATS Lasso) to penalize groups formed by all coefficients that belong to the same individual-specific variable in MNLs, we introduce for the first time a tailor made adjustment of the Lasso that explicitly incorporates the alternative-wise specification of choice-specific covariates, and offer an in-depth discussion of the resulting variable selection and its practical implications in the context of party choice modeling.

[4] While the alternative-wise specification of choice-specific attributes is widely used and applied in transportation economics and econometrics (see, e.g., Ben-Akiva and Lerman, 1985; Hensher et al., 2007; Louviere et al., 2009; Train, 2009), it has received little attention in the empirical study of voting behavior. See as an early exception (Thurner, 2000), and the more detailed recent illustration by Mauerer et al. (2015).

[5] All codes and routines to implement the proposed model will be made available on the authors' website.

[6] Since Downs' seminal work on democracy as a political market (Downs, 1957), a growing number of scholars theoretically as well as technically developed this approach (see, e.g., Alvarez and Nagler, 1998; Thurner, 2000; Dow and Endersby, 2004; Adams et al., 2005; Schofield et al., 1998). For a summary of the formal set-up of the Spatial Theory of Voting, see Thurner (2000).

[7] Note that we explicitly use respondent-specific perceptions of party positions instead of perceived mean party placements. By using subjective self-placements of party positions, we follow previous studies who offer both theoretical arguments and empirical evidence in favor of voter-specific