ARTYKUŁY

ASK, 2007, nr 16, strony 9–23
Copyright by ASK, ISSN 1234–9224

Joshua Kjerulf Dubrow
Polish Academy of Sciences

CHOOSING AMONG DISCRETE CHOICE MODELS FOR VOTING BEHAVIOR*

Analyses presented in this paper aim at testing demographic cues hypothesis, which explains voting behavior as a function of the distance between the voter and the object of the vote, expressed as demographic similarity. Four types of multivariate regression models – binomial logistic (BNL), multinomial logistic (MNL), contrast logistic (CONTRAST), and conditional logistic (CLOGIT) – are applied to explain vote choice among Polish parties in the 1997 parliamentary election. For all models I use survey data combined with information on political parties derived from characteristics of the electoral candidates. The results demonstrate that for testing demographic cues hypothesis CLOGIT and BNL are the most advisable options in terms of elucidation of the regression coefficients; MNL and CONTRAST involve cumbersome interpretation and their fit to the theory is questionable.
Key words: logistic regression, demographic cues hypothesis, discrete choice models, voting behavior

Researchers are often confronted with the problems of explaining choices. Among these problems is explaining voting behavior in democratic elections, particularly that of individuals' preferences for political parties. Which models are the most appropriate for accounting for party preferences? This paper addresses statistical modeling where the outcome variable is a discrete choice among political parties. I explore four types of multivariate regression models: binomial logistic (BNL), multinomial logistic (MNL), contrast logistic (CONTRAST), and conditional logistic (CLOGIT). To illustrate an application of

* The author gratefully acknowledges Kazimierz M. Slomczynski, Pamela Paxton, and Irina Tomescu-Dubrow for their helpful comments and suggestions. Please direct all correspondence to Joshua Kjerulf Dubrow, Institute of Philosophy and Sociology, Polish Academy of Sciences, Nowy Swiat 72, 00-330 Warszawa, Poland; e-mail: jdubrow@ifispan.waw.pl.

these models, I provide empirical examples of vote choice among Polish parties in 1997 as the outcome variable.

What is Discrete Choice?

A discrete choice is any preference selected from a set of independent alternatives. If among objects a, b, and c the individual chooses b, then this is the individual's discrete choice; on a scale of 1 or 0, choice b becomes 1 and alternatives a and c become 0. Such choice contrasts with standard models in which the quantity of each object is assumed to be continuous variable.

A discrete choice model is an econometric model that assumes choices to be substantially independent of one another. Most discrete choice models impose the Irrelevance of Independent Alternatives (IIA) assumption. In explaining vote choice, for example, IIA "implies that in a contest between a liberal and a conservative party, the entry of a second conservative party would not alter the relative probability of an individual voter choosing between the two initial parties. However... the two conservative parties are close together in issue space and hence are likely to be viewed as substitutes by voters..." (Alvarez and Nagler 1997: 57).

In models of voting, determining exactly when the IIA assumption is violated depends on which "issue spaces" are relevant. For example, in 1997, Poles faced four choices among the major (based on the percentage of popular vote) parties: Sojusz Lewicy Demokratycznej (SLD), a post-communist leftist party; Polskie Stronnictwo Ludowe (PSL), a farmers' interests party; Unia Wolności (UW), a rightist party; and Akcja Wyborcza Solidarność (AWS), a rightist, religiously oriented party comprised of the splintered Solidarity movement parties. These choices are discrete depending on how the issue space is defined. Taking only economic orientation into account, a discrete choice model with SLD and PSL as alternatives violates the IIA assumption, as SLD and PSL are similarly situated in terms of economic issues (in 1997, both straddled the line between "statism" and economic liberalism.) Adding the religious dimension – PSL leans toward Catholic traditionalism and SLD leans toward anti-clericalism – changes what constitutes the total issue space. Thus, there are also grounds to suspect that these parties are sufficiently dissimilar such that IIA is not a critical issue.[1]

[1] When IIA is a problematic assumption, some advocate a fifth discrete choice model: multinomial probit (MNP) (Alvarez and Nagler 1997). Multinomial probit is similar to MNL and has the added bonus of relaxing the IIA assumption inherent in MNL and BNL. However,

Demographic Cues Hypothesis

Analyses presented in this paper aim at testing demographic cues hypothesis, which explains voting behavior as a function of the distance between the voter and the object of the vote, expressed as demographic similarity. Assume that the voter has demographic characteristic V_x and that the potential alternative that he or she is taking into account is characterized by the same demographic characteristic O_x. According to the demographic cues hypothesis, the ultimate choice C is a function of V_x, O_x, and the similarity between V_x and O_x. The hypothesis states that given V_x and O_x, the probability of choosing C increases if V_x and O_x are closer.

Most analyses exploring demographic cues hypothesis posit individual candidates as the object of the vote (Cutler 2002; McDermott 1997; Sanbonmatsu 2004). In parliamentary democracies such as Poland, voters can vote for candidates or whole party lists. In the models that follow, I measure the object of the vote as parties, and not as individual candidates. Following the spatial model of voting, the demographic cues hypothesis is that *voters tend to vote for the party whose demographic composition is the most demographically similar to them*. For example, women will tend to vote for parties that are the most “women friendly,” which I measure as being the party that has the most women candidates. Political theory supports the conjecture that members of disadvantaged groups are likely to vote for parties having on their electoral lists a relatively large proportion of candidates with the same demographic characteristics (see Dubrow, forthcoming).

In the analyses presented in this paper I consider only votes for major parties in the 1997 elections: SLD, PSL, UW, and AWS. Assume that we are interested in voting for SLD. Then, the generic model for testing the demographic cues hypothesis is:

$$\log (p/ 1 - p) = a + b_1G + b_2X + b_3Z$$

where p denotes probability of voting for SLD, G refers to voter gender (1 = female, 0 = male), X is a dummy variable describing the position of SLD with regard to the proportion of female candidates in a given district (1 = the highest proportion among all parties, 0 = otherwise), and Z is the interaction term of G and X (Z = G*X). According to the demographic cues hypothesis one would

MNP produces coefficients that are not-intuitive and, hence, difficult to interpret and has been shown to not have any other particular advantages over MNL (see Dow & Endersby 2004).

expect that women vote for SLD if this party in a given district is leading in the proportion of female candidates, i.e., $b_1 + b_3 > 0$.

Data

Data on individual voters and their vote decisions are contained in POLPAN (see Slomczynski 2002), a panel dataset representative of Poles, first collected in 1988 and continued every five years thereafter. Party characteristics used to derive the most demographically similar party are from POLCAN, a universe of Polish candidates for every post-communist election to date (Zielinski, Shabad, and Slomczynski 2005).

To attach the party characteristics to individual voters, I merged POLPAN with POLCAN. Observations in POLPAN and POLCAN can be aggregated at the administrative district level, or voivodeship[2]. I merged the two datasets by (a) computing the percentage of women candidates per party in all voivodeships, and (b) appending a derived score to each individual voter within their voivodeship. For example, in 1997 in Warsaw, SLD had 18% women candidates, PSL also 18%, UW 22% and AWS 8%. I computed separate variables for each party and assigned the derived scores to individual voters in Warsaw, such that all voters in that district obtain one score for SLD, one for PSL, one for UW, and one for AWS. In each of 49 voivodeships, the party with the most women candidates receives a score of 1, while all other parties receive a score of 0. This variable reflects a competition among the top parties on the voivodeship level. Thus, I assume that voters compare parties in their districts.

In testing the hypothesis, interaction terms are the key independent variables as they measure the fit of voter-party demographics. To compute interaction terms, I multiplied the voter characteristic and the party characteristic. For example, to derive a voter-party fit based on the gender of the voter and the gender composition of the party, the interaction term has a value of 1 if the voter is a woman in a district in which the given party leads all other parties in the proportion of women candidates; otherwise, the value is 0.

For the analyses that follow, I focus on the election of 1997 and each party's gender demographic composition – computed from the proportion of women candidates – for the top four parties in that election: SLD, PSL, UW, and AWS.

[2] Voivodeships were not strictly comparable across data sets. For an explanation of how this was reconciled, see Dubrow 2006.

Cases are restricted to only those who voted in 1997. To simplify basic illustrations, in all models I do not include standard voting behavior variables such as age and social class of the respondent or attitudes toward the economy (for models with controls that test demographic cues theory, see Dubrow 2006). Analyses for BNL, MNL, and CONTRAST were performed using SPSS and analyses for CLOGIT were performed using STATA.

General Properties of the Models

Table 1 outlines general properties of the four discrete choice models. The main conceptual difference between these models pertains to the outcome variable. In a binomial logistic regression, BNL, a probability of voting for a given party (p_i) is contrasted with non-voting for that party ($1 - p_i$), that, for active voters, means a preference for other parties; the estimated function is of logistic form: $\log [p_i/(1 - p_i)]$. In the case of multinomial logistic regression, MNL, the odds of voting for each party from the subset of all parties is compared with a reference unit constituted by the party not included in the subset. Usually, the subset contains all parties but one and that remaining party is a reference unit. Contrast logistic regression, CONTRAST, is a binomial logistic regression performed on pairs of parties. Finally, in the conditional logistic regression, CLOGIT, the outcome variable is a choice of a party among all alternatives within a set of person-choice observations.

Units of analysis differ by model. In BNL, the individual voter is the unit of analysis and in some models the observations in the reference point may also include non-voters. In MNL, only voters of parties present in the outcome variable are considered. In CONTRAST, the units of analysis are limited to those who voted for either of the party pairs. CLOGIT follows a different logic. In a binary setting, one can assume that each voter separately decides whether to vote for party a, b, or c. Voting for a, b, and c constitutes distinguishable trials. For this reason in CLOGIT – in contrast to BNL, MNL and CONTRAST – the unit of analysis is voter-party, or person-choice.

Voters' characteristics are present in all models and in the same form for BNL, MNL, and CONTRAST. In CLOGIT, voter characteristics must vary within grouped units of person-choices; thus, voter's characteristics take the form of interactions between voter's demographics and a party value assigned by the researcher.

Theory and peculiarities in data arrangement determine the levels at which parties' characteristics are aggregated. Since demographic cues theory does not strictly posit the level of aggregation in which voters assess the demographic composition of parties, following the adage that "all politics is local," for most analyses I assume that voters are most capable of assessing the demographic composition of the parties within their voivodeships. In every model except CLOGIT, party characteristics are aggregated at the voivodeship level. Owing to the data arrangement typical for conducting CLOGIT, party characteristics can be aggregated not only at the voivodeship level but also represented as emergent properties. In order to test the hypothesis, a higher level of analysis – that is, across voivodeships, or on "national scale" – is introduced.

Table 1. General Properties of Discrete Choice Models That Explain Vote Behavior

Discrete Choice Models	Vote Choice	Unit of Analysis	Voters' Characteristics Present?	Parties' Characteristics
BNL – Binomial logistic	Given party versus the rest	Voter (potentially also non-voter)	Yes	Voivodeship level – only relevant party
MNL – Multinomial logistic	Each party with a reference point	Voter (all parties)	Yes	Voivodeship level – all parties
CONTRAST Contrast logistic [a]	One party against another	Voter (for compared parties only)	Yes	Voivodeship level – compared parties
CLOGIT – Conditional logistic	Party among alternatives[b]	Voter-party, or person-choice (all parties)	Yes – as interactions	Both Voivodeship and national level – all parties

[a] Binomial logistic for selected pairs of choices.
[b] Since the unit of analysis is voter-party, the variable "party choice" does not distinguish between individual parties in the same manner as the other models; rather, the vote choice variable is derived from the characteristics of the parties considered.

Binomial and Multinomial Logistic Regression, BNL and MNL

Binomial logistic regression (BNL) is the simplest statistical tool that can test the hypothesis. Generally, in testing this hypothesis by BNL, the choice variable is a given party versus all others. In the illustrative case presented in the first part of Table 2, vote for SLD = 1, otherwise = 0. The unit of analysis is the individual voter who is characterized by gender. Party characteristic is limited to whether SLD had the most women candidates in their party compared to all other top parties within a given voivodeship. The interaction term is computed with the voter and party characteristics, such that women who live in a voivodeship where SLD has the most women candidates in their party in comparison to all other top parties = 1, otherwise = 0.

The interaction term is negative but not significant. Generally, according to the results presented in the first part of Table 2 it is not true that women in districts where SLD has the most women candidates tend to vote for SLD. When SLD has the most women candidates among all parties, voters are more likely to vote for SLD independently of gender. Thus, in view of results in Table 2 for SLD, the demographic cues hypothesis should be rejected.

Table 2. Binomial Logistic Regression (BNL) of Vote for SLD, PSL, and UW on Voter's Gender, Party Composition, and Interaction Term, 1997

	B	SE	Exp(B)
Voting for SLD[a]			
Voter's gender, G (1 = female, 0 = male)	0.03	0.15	1.03
Party (SLD) leading in the proportion of female candidates, X (1 = yes, 0 = else)	0.57**	0.19	1.77
Interaction of G and X, Z (Z = G * X)	-0.14	0.28	0.87
Constant	-0.80***	0.11	0.45
Voting for PSL[b]			
Voter's gender, G (1 = female, 0 = male)	-0.06	0.21	0.94
Party (PSL) leading in the proportion of female candidates, X (1 = yes, 0 = else)	-0.08	0.38	0.92
Interaction of G and X, Z (Z = G * X)	-0.42	0.59	0.66
Constant	-2.08***	0.14	0.13
Voting for UW[c]			
Voter's gender, G (1 = female, 0 = male)	-0.25	0.26	0.78
Party (UW) leading in the proportion of female candidates, X (1 = yes, 0 = else)	0.07	0.23	1.08
Interaction of G and X, Z (Z = G * X)	0.35	0.34	1.41
Constant	-1.70***	0.17	0.18

[a] Chi Square = 13.20 (df = 3), Cox/Snell R^2 = 0.012
[b] Chi Square = 1.81 (df = 3), Cox/Snell R^2 = 0.002
[c] Chi Square = 3.11 (df = 3), Cox/Snell R^2 = 0.003
*** $p < 0.001$ ** $p < 0.01$ * $p < .05$

In contrast with the case of SLD, when PSL has the most women candidates among all parties, voters are not likely to vote for PSL. Similarly, when UW has the most women candidates among all parties, voters are not likely to vote for UW. In the cases of PSL and UW, neither gender alone nor the interaction term of gender and appropriate party characteristics have statistically significant effects. Note, however, that in the case of UW the interaction term is large. It is not significant due to almost equally large standard error.

To model discrete choice, especially in voting, some researchers use MNL. However, Alvarez and Nagler (1998) argue that in discrete choice models, MNL and BNL posit "the same choice processes" (64). When the authors conducted a vote choice model, "ocular examination" revealed "that they produce consistent estimates of the same parameters" (64) – meaning that the models look the same. Thus, MNL may have no particular advantages over BNL.

Table 3. Multinomial Logistic Regression (MNL) of Voting for SLD, PSL, and UW (in Comparison with Voting for AWS) on Voter's Gender and Parties' Role in Supporting Female Candidates, 1997

	SLD		PSL		UW	
	B	SE	B	SE	B	SE
Gender, G (1 = female, 0 = male)	0.25	0.49	1.29†	0.77	-0.08	0.58
SLD leading in the proportion of female candidates, X_{SLD}	-0.15	0.36	0.93†	0.53	-1.31**	0.46
Interaction, Z (Z = G * X_{SLD})	-0.23	0.54	-0.63	0.80	0.17	0.70
PSL leading in the proportion of female candidates, X_{PSL}	-0.38	0.39	0.82	0.53	-0.74	0.47
Interaction, Z (Z = G * X_{PSL})	-0.33	0.55	-1.50†	0.83	-0.50	0.70
UW leading in the proportion of female candidates, X_{UW}	-0.72*	0.34	1.36**	0.50	-0.91*	0.40
Interaction, Z (Z = G * X_{UW})	-0.29	0.50	-1.69*	0.76	0.11	0.60
Constant	0.37	0.33	-2.48***	0.52	0.05	0.38

Chi Square = 59.10 (df = 21), Cox/Snell R^2 = 0.05
*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ † $p < 0.10$

Because I test a hypothesis that necessarily includes variables specific to particular parties, and as existing statistical software for MNL forces all party characteristic variables into the same equation, MNL is a suboptimal choice. The following illustrates MNL. Categories of the outcome variable are the top parties of 1997: SLD = 1, PSL = 2, and UW = 3. The choice variable is each party, with AWS

= 4 as the reference point. The unit of analysis is individual voter who voted for one of these parties in the 1997 election. I included the voter and party characteristic variables and the interaction terms as in the BNL model above.

In the cases of SLD and UW, the sum of the coefficients for gender and the interaction term for these parties is positive but not significant for $p < 0.05$. Only for UW the interaction term is positive with the Exp(B) value = 1.11 (but not significant). In the case of PSL the corresponding sum is negative, which is contradictory to the sign implied by the demographic cues hypothesis. Generally, in statistical terms, the hypothesis is rejected, mainly due to the large standard errors.

Analysis of the standard errors and their impact on the decision about rejecting the hypothesis is an important part of statistical analysis. SPSS program provides 95% confidence intervals for Exp(B), that is odds ratios. If we take upper bounds of the Exp(B) for each party, the picture is different. Here are upper bound values of Exp(X) for each party:

SLD – Gender 3.34, SLD leading in the proportion of female candidates 1.75, and Interaction term 2.26;

PSL – Gender 16.32, PSL leading in the proportion of female candidates 6.43, and Interaction term 1.13;

UW – Gender 2.88, UW leading in the proportion of female candidates 0.88, and Interaction term 3.64.

For SLD and PSL the proportion of female candidates in electoral districts positively influences the probability of voting for these parties, independently of gender. However, if we compare women to men we would notice a clear gender difference in voting behavior. For all three parties, women in districts in which a given party leads in the proportion of female candidates have higher propensity to vote for this particular party than men do. The difference is very substantial since odds ratios for gender and for interaction terms are high (from 2.88 to 16.32 for gender and from 1.13 to 3.64 for the interaction term).

In the SLD model, PSL's and UW's characteristics have no meaningful interpretation and are possibly interfering in a non-theoretical way with the variables that are of specific interest, i.e. the SLD variables. To my knowledge, there is no statistical software that would enable the researcher to effectively "block out" the party characteristics of the other parties in MNL. Thus, although MNL allows for explicit inter-party comparison of determinants of voting preferences, it is ill-suited to the task at hand.

Contrast models, CONTRAST

Contrast models are an option, but in addressing the hypothesis, they are not a substantial improvement over BNL. To conduct CONTRAST, I created pairs out of the top parties, truncating the sample to only those who voted for either party in each pair. I then employed a series of BNL regressions on the party pairs.

To demonstrate the main features of CONTRAST, I performed a BNL regression on one of the possible subsets of top parties in 1997. The choice variable is one party versus another; in this case, SLD = 1, AWS = 0. The unit of analysis is the individual voter who voted for either SLD or AWS in the 1997 election; all other cases are eliminated. Voter characteristic is limited to gender, where female = 1. Party characteristic is limited to whether SLD has more women than AWS in a given voivodeship. The interaction term is computed with the voter and party characteristics, such that women who live in a voivodeship where SLD has more women than AWS = 1, otherwise = 0.

Table 4 presents this simple model. In this model, only the party characteristic variable is significant, suggesting that when SLD has more women in their party than AWS, voters of either gender are more likely to vote for SLD. Thus, in view of results in the table, the demographic cues hypothesis should be rejected. This result corresponds to that presented in the first part of Table 2.

Table 4. Contrast Model (CONTRAST) of Voting for SLD or AWS on Voter's Gender, Party Composition, and Interaction Term, 1997

	SLD =1, AWS = 0		
	b	SE	Exp(B)
Voter's gender (female = 1)	-0.06	0.16	0.94
SLD leading AWS in the proportion of female candidates (yes =1, else = 0)	0.42*	0.21	1.52
Interaction of Voter * Party	0.07	0.31	1.07
Constant	-0.23*	0.12	0.79

Chi Square = 8.71 (df = 3), Cox/Snell R^2 = 0.010
* $p < 0.05$

To test demographic cues hypothesis, CONTRAST seems a suboptimal option. CONTRAST is acceptable only under the assumption that voters choose between two parties, rather than all relevant options which, for 1997 at least, consist of four different parties. There is no reason, *a priori*, to assume that voters do this, e.g. that SLD voters are only choosing between SLD and AWS.

Most likely, they evaluate all major parties and choose the one that best suits them. Statistically, the CONTRAST model can be estimated, but theoretically, it is counter-intuitive for testing demographic cues hypothesis.

Conditional Logistic Regression, CLOGIT

Alvarez and Nagler (1997: 56) define CLOGIT as a regression model that is "conditional on the characteristics of the choices; thus, it explicitly allows for measures of party characteristics." The actual CLOGIT equation can be found in McFadden (1974) and Long (1997). CLOGIT is similar to MNL (Long 1997), with some key differences explained below. Like MNL, CLOGIT assumes IIA.

I performed CLOGIT analysis using STATA statistical software[3]. To simplify the analysis, I focus on the top three parties of 1997 – SLD, PSL, and AWS, using voter's gender and a party characteristic that refers to the rank according to the proportion of women candidates. I constructed a hypothetical dataset, with AWS, PSL, and SLD as the parties (Table 5). Case identification codes (ID) are aligned such that three consecutive units are paired with the three choice possibilities; thus, the unit of analysis becomes person-choice. Choice among the three parties varies within each ID group; thus, the dichotomous outcome variable, or vote choice, is a given party among alternatives.

There are two covariates for the choice: the first is chooser-specific (voter characteristic) and the second is choice-specific (party characteristic). Here, the voter characteristic is the respondent's gender. Party characteristic refers to the proportion of female candidates. Parties are ranked from 1 to 3, with 1 referring to the party with the fewest female candidates.

In Table 5, hypothetical voter ID 1 (male) voted for PSL as PSL has the second most women candidates, whereas ID 3 (female) voted for SLD as SLD has the most women candidates.

As for the actual data analysis, I begin by transforming individual voters into person-choices as the units of analysis and by creating a new dependent variable that is suited to the CLOGIT data arrangement. For CLOGIT models, data arrangement requires that for each individual – represented by a group of the same IDs – the outcome variable is the choice of the object (in this case, a party). The new dichotomous variable reflects vote choice as conditional on the

[3] The advantage of STATA is that it has a pre-set command to run CLOGIT (also named CLOGIT), whereas SPSS only has a proxy procedure.

characteristics of the voter and the objects (parties) presented within a given group of IDs.

Table 5. Hypothetical CLOGIT Data Arrangement

Voter's ID	Party	Vote Choice	Voter Characteristic	Party Characteristic
1	AWS	0	male	1
1	PSL	1	male	2
1	SLD	0	male	3
2	AWS	1	female	1
2	PSL	0	female	2
2	SLD	0	female	3
3	AWS	0	female	1
3	PSL	0	female	2
3	SLD	1	female	3
...	...	...	...	...

Since the unit of analysis is person-choice, all voter characteristics must vary *within* each group of ID's *and* by party. To obtain chooser-specific effects I construct interactions between voter's gender and an assigned party value. Specifically, voter's gender is multiplied by the party value for SLD = 3, for PSL = 2, and for AWS = 1.

There are two types of party characteristics. One type is constructed from values at the voivodeship level. The party value (1/0) – whether, at the voivodeship level, the party leads in proportion of female candidates in comparison to all other top parties – is multiplied by voter's gender. The interaction terms are computed for SLD and PSL. The meaning of the first variable is "women in voivodeship in which SLD has the highest proportion of female candidates," and the meaning of the second variable is "women in voivodeship in which PSL has the highest proportion of female candidates."

The second choice-specific variable refers to a non-voivodeship level variable. I ranked SLD, PSL, and AWS according to their proportion of women candidates aggregated across districts. Nationally, in 1997, SLD had the most women candidates, receiving the highest rank (3); PSL had the second most (2) and AWS ranked lowest (1). In this illustration I test the assumption that voters base their party preferences on both the district *and* national level image of each party, measured by the descriptive representation of women in party candidate lists.

Table 6. Fixed-Effects Conditional Logistic Regression (CLOGIT) of Party Choice on Voter Characteristic, Voivodeship Level Party Characteristics, and National Party Image

	B	SE	z	P>|z|	95% Confidence Interval	
Voter's gender	0.13	0.09	1.40	0.16	-0.32	0.05
Women in districts where SLD is leading in the proportion of female candidates	0.32	0.13	2.39	0.02	0.06	0.59
Women in districts where PSL is leading in the proportion of female candidates	0.01	0.16	0.07	0.95	-0.30	0.32
National party image[a]	-0.07	0.06	-1.22	0.22	-0.18	0.04

Number of observations = 2868, LR Chi^2 / df (4) = 11.55, Prob > Chi^2 = 0.02
Log likelihood = -1044.5
[a] Parties ordered according to the proportion of female candidates.

Table 6 presents the model. Model fit is satisfactory. The most important result is that the voivodeship level party characteristic variable is statistically significant for SLD, indicating that when SLD has the most women candidates in a given voivodeship, women voters are more likely to vote for that party. When PSL has the most women candidates in a given voivodeship, women voters are also more likely to vote for PSL, although the result is not significant. This model suggests that in the absence of standard voting behavior controls, voters choose parties disregarding their national image.

CLOGIT is helpful for testing the demographic cues hypothesis. As for advantages, CLOGIT allows the researcher to test whether taking into consideration the choice among the characteristic of the parties variable matters. Specifically, CLOGIT provides a more detailed test of the theory, pitting two different theoretical assumptions – whether voters consider voivodeship image and/or national image – in the same model. The main disadvantage is that the effects of all individual and voivodeship level characteristics are expressed in terms of interactions.

Summary and Discussion

In this paper I explored four discrete choice models of voting behavior to test an elementary form of the demographic cues hypothesis: BNL, MNL, CONTRAST, and CLOGIT. By providing basic versions of these models, without standard voting behavior controls, I was able to demonstrate the general properties of these models along with the advantages and disadvantages of each. In terms of capability to test the hypothesis, CLOGIT and BNL are the most advisable while MNL and CONTRAST are the least advisable.

CLOGIT's main advantage is that it explicitly allows for choice to be conditional on the demographic characteristics of the parties. Simplicity is BNL's greatest advantage. Although BNL aggregates non-chosen parties into the reference point, it has been shown in other research to produce effects comparable to that of MNL. However, MNL is not advisable option as it forces non-theoretically relevant explanatory variables into the model. CONTRAST is suboptimal since it forces the user to construct theoretically counter-intuitive outcome variables.

Note, too, that the CLOGIT and BNL models differ in their empirical support for the hypothesis. While a full examination of the reasons why there is a difference is beyond the scope of this paper, one factor is important to stress: CLOGIT allows vote choice to be conditional on the characteristics of the party.

This exploration into discrete choice models serves also as a reminder that model choice should be based on the research question and the theory behind it. Take, for example, a demographic cues hypothesis where presidential candidates are the object of the vote. In a line-up of three presidential candidates, voters base their choice on whether the candidate is the most demographically similar to them (e.g. *ceteris paribus*, women voters prefer female candidates for president). Here, CLOGIT would be the best option as vote choice is conditional on the demographic characteristics of the candidates. Moreover, since this vote choice occurs only at one level – national – the peculiarities of CLOGIT's data arrangement align perfectly with the hypothesis. BNL would be the next best option. It produces similar effects to MNL but, unlike MNL, doe not force unnecessary variables into the model. CONTRAST would still rank last, unless the theory specifically assumes that a certain type of voter would choose only between two of the three candidates. In that case, CONTRAST would be as good an option as BNL. Thus, in choosing a discrete choice model, theory should be the primary consideration and the researcher's most reliable guide.

References

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

Cutler, Fred. 2002. "The Simplest Shortcut of All: Sociodemographic Characteristics and Electoral Choice." *Journal of Politics* 64(2): 466–90.

Dow, Jay K. and James W. Endersby. 2004. "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research." *Electoral Studies* 23(1): 107–22.

Dubrow, Joshua Kjerulf. 2006. *Enhancing Descriptive Representation in a New Democracy: A Political Market Approach*. PhD dissertation, The Ohio State University, Columbus, Ohio, USA.

Dubrow, Joshua Kjerulf. "Voting for Descriptive Representation: Demographic Cues across Multiple Elections, 1993–2001." Forthcoming in *Continuity and Change in Social Life: Structural and Psychological Adjustment in Poland*, edited by Kazimierz M. Slomczynski and Sandra Marquart-Pyatt. IFiS: Warsaw.

Long, Scott J. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

McDermott, Monika L. 1997. "Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary United States Elections." *American Journal of Political Science* 41(1):270–283.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." Pp. 105–42 in *Frontiers in Econometrics*, edited by Paul Zarembka. New York: Academic Press.

Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46(1): 20–34.

Slomczynski, Kazimierz M. 2002. *Social Structure: Changes and Linkages*. Warsaw: IFiS Publishers.

Zielinski, Jakub, Kazimierz M. Slomczynski, and Goldie Shabad. 2005. "Electoral Control in New Democracies: The Perverse Incentives of Fluid Party Systems." *World Politics* 57: 365–95.