Department of Political Science, University of Aarhus Tel.: 89 42 12 71
Bartholins Allé, 8000 Aarhus C Fax.: 86 13 98 39
Søren Risbjerg Thomsen E-mail:srt@ps.au.dk

17 November 2011

Revised version of a paper prepared for presentation to the 2009 Annual Meeting of the Midwest Political Science Association, Palmer House Hilton, Chicago, Illinois, April 2-5, 2009.

The cultural component in voting behaviour[1]

Why is it so simple?

In most democracies it is very common observation on election nights that parties with a clear advance, compared to the previous election, in the first scattered results continue with this trend until the final results are counted. Similar, parties with decline in the first results often tend to continue this trend to the end. This simple pattern of "uniform swing" is the very reason that one can often make quite good election forecasts from the first few results in national elections (Butler and van Beek, 1990; Karandikar et al., 2002).

But why is this simple pattern of uniform swing across election districts so common? Since parties tend to represent different special interests, on should rather expect that a party becoming more competent in pursuing its policies, should advance in election districts with high shares of voters agreeing with its policies, and decline elsewhere. As pointed out by Butler and Stokes (1969) an obvious explanation is that politics is not only about issues dividing the electorate but also about the so called "valence issues" where everybody thinks that "more is better". Valence issues are often about solving current economic problems or about the more general competence of the current government in solving problems facing the nation as a whole. In this situation the general consensus on the competence of the individual party leaders in government as well as in the opposition is often paramount.

The importance of valence issues as witnessed by the prevalence of uniform swing is an indication that common values and perceptions are more important than rational choice theory would like us to think. Defining culture as shared values (Inglehart 1997 p. 15) the salience of valence issues can be understood as an important cultural component in voting behaviour.

As a first step in an effort to build a comprehensive model for both individual and aggregate-level voting behaviour this paper searches for an individual-level model that is able to separate the effect of valence issues common to all voters from the effect of so called "position issues" that are evaluated different by the individual voters.

Discrete choice models for individual-level voting behaviour

Since the 1990's probability models for discrete choice borrowed from econometrics have been widely employed in quantitative studies of individual level voting behaviour. The problem is to build a model that can use available data to explain the probability P_{ij} that voter i chooses party or candidate j.[2] Such a model can formally be derived by assuming that the utility U_{ij} of party j for voter i can be partitioned in a fixed part u_{ij} and a stochastic error component ε_{ij}.

[1] Thanks to the national social science data archives in Denmark, Norway, Sweden and UK, The Irish National Election Study, and the ICPSR in the US for making election surveys available to the author.

[2] To make it simple, in the following we will only talk about party choice, as in parliamentary elections, but it could as well be choice of a candidate, as in presidential elections.

(1) $$U_{ij} = u_{ij} + \varepsilon_{ij}\ ; \quad i = 1,\ldots,n\ ; \quad j = 1,\ldots,J\ .$$

Where *n* is the number of respondents in a voter survey and *J* is the number of parties. The idea is that the voter when confronted with a choice situation does not only have a fixed utility associated with a particular party, but for each party additional is making a random draw from a distribution generated by the distribution of ε_{ij}. The choice probability becomes particular simple by assuming that the errors are independent, identical distributed (iid) with a type I extreme value distribution (Dow and Endersby, 2004). In this case the probability that voter *i* chooses party *j* becomes the multiple logit model

(2) $$P_{ij} = \frac{e^{u_{ij}}}{\sum_{h=1}^{J} e^{u_{ih}}}$$

Thus, the probability is determined by setting the exponential transformation of the fixed utility in relation to the sum across all parties of this transformed utility.

A much criticized property of this multiple choice model in equation (2) is the assumption of "Independence of Irrelevant Alternatives" (IIA) i.e., the assumption that the ratio between the probabilities of two different parties only depends on the fixed utilities of these two parties, since it can be derived from equation (2) that

(3) $$\frac{P_{i1}}{P_{i2}} = \frac{e^{u_{i1}}}{e^{u_{i2}}}.$$

An example that points to the absurdity of this property is the Red Bus/Blue Bus example. If a commuter initial only has the choice between using a car or a red bus with a certain probability ratio between the two modes of transportation, and a blue bus then comes along that goes to the same destination for the same price, surely the original ratio will be affected. Similar, if one more left wing party runs in an electoral district, this might decrease the support for other left wing parties. However, this might not be a problem in studies of voting behaviour, at least not in proportional systems where all parties run in all constituencies. It is more likely a problem in transportation research, where different commuters have different sets of travel alternatives.

Troubled by the IIA problem another and more flexible assumption about the utility errors was suggested by Alvarez and Nagler (1998). By this assumption it is instead assumed that the errors are distributed multivariate normal, with zero means and with different covariances to be estimated between the errors. This leads to the so called Multinomial Probit Model (MPM). However, recent research has pointed out that with appropriate specification of equation (2) the IIA problem might not be so serious, and with the number of observations about 1,000-2,000 in electoral studies the multinomial probit model has serious estimation problems (Dow and Endersby, 2004; Long and Freese, 2006, p. 322). For these reasons we will in the rest of this paper only focus on models derived from the simpler multiple logit model in equation (2). Since we have "alternative-specific" data for each party we use the Conditional Logit Model (CL) (Long, 1999, pp. 178-182) to estimate the different models for voting behaviour.

Separation of the effect of valence and position issues

A simple way to separate the effects of on the one hand valence issues and on the other hand position issues is the following partitioning of the fixed utility in equation (2) in two components

$$u_{ij} = a_j + b_{ij} \quad (4)$$

where a_j is the effect of valence issues only depending on the party j and b_{ij} is the effect of position issues depending on the combination of the positions of the voter and the position of the party on these issues. While b_{ij} is usually a function of empirical indicators, a_j is estimated indirectly as a party-specific intercept indicating the "general popularity" of the party caused by valence issues. The party intercept (or "party constant") is often omitted in much work on issue voting in an effort to explain party support as caused only by positions in a policy space (see for example Merrill et al. 1999, Adams et al. 2005), but this might lead to arbitrary dependence on the scale of the independent variables. The work of Schofield and Sened (2006) is a prominent examples of the use of party intercepts expressing the effect of valence issues.

A useful property of the multiple logit model (2) is that any idiosyncratic additive property of the individual voter is cancelled. Assume for example that every voter has an idiosyncratic bias z_i across all parties for making especially positive or especially negative utility evaluations, then one should think that equation (4) should be changed to

$$u_{ij} = a_j + b_{ij} + z_i \ . \quad (5)$$

However, when inserting eq. (5) into eq. (2) z_i is cancelled:

$$P_{ij} = \frac{e^{u_{ij}}}{\sum_{h=1}^{J} e^{u_{ih}}} = \frac{e^{a_j + b_{ij} + z_i}}{\sum_{h=1}^{J} e^{a_h + b_{ih} + z_i}} = \frac{e^{z_i} e^{a_j + b_{ij}}}{e^{z_i} \sum_{h=1}^{J} e^{a_h + b_{ih}}} = \frac{e^{a_j + b_{ij}}}{\sum_{h=1}^{J} e^{a_h + b_{ih}}}. \quad (6)$$

Thus, when applying the multiple logit model (2) there is no difference between equation (4) and (5) and one can cancel any fixed effect only dependent on i.

Proximity versus direction

Since the general popularity a_j in equation (4) is measured indirectly as an intercept, the actual measurement of a_j is highly dependent on how the effect b_{ij} of position issues is measured.

Assume for the moment that voters as well as parties only vary in position along a single left-right dimension. X_i indicates the position of voter i and Y_j indicates the position of party j on this dimension. According to the work of Downs (1951) the closer the party is to the voter the higher is the utility. A very common proximity model is

$$b_{ij} = -(X_i - Y_j)^2 . \quad (7)$$

With this model, the more different the party is from the voter the less the voter tends to vote for the party. Since utility is a relative measure it is no problem that b_{ij} with the proximity model (7) is a non-positive value.

The alternative directional model is based on the assumption that the utility of a party is higher, the more extreme the party is in the same direction as the voter. The idea is that the

policy signal is stronger for the more extreme parties and that the voter can benefit from supporting such parties since likely political compromises are drawn closer to the voter's own position. Assuming that the neutral middle point of the left-right scale (also called the status quo point) is 0, then the model for the effect of position issues is

$$b_{ij} = X_i Y_j . \tag{8}$$

With this model the utility of a party is higher the more extreme both X_i and Y_j are with the same sign (positive or negative). According to the literature there should be a limit or "penalty" for very extreme parties (Rabinowitz and Macdonald 1989, p. 108), but this condition is rarely included in model applications.

For multiple issues the proposed proximity model with K issues is

$$u_{ij}^{p} = a_j^{p} - \sum_{k=1}^{K} c_k^{p} (X_{ki} - Y_{kj})^2 ; \quad i = 1,\ldots,n ; \quad j = 1,\ldots,J , \tag{9}$$

where the "exponent" p is not an exponent but an indication that the parameter belongs to the proximity model. X_{ki} is the self-placement of voter i on issue scale no. k and Y_{kj} is the position of party j on the same scale measured by the mean voter assessment. The more different X_{ki} is from Y_{kj}, the less utility of the party. Some researchers prefer to use the individual voter assessment of the party position for modelling the utility for each voter, but this raises the problem of reverse causality i.e., the voter assess the position of a party she favours close to her own position. A further argument for using mean voter assessment is that the party position is a national-level attribute mostly decided by national politics in national elections. c_k^{p} is a coefficient measuring the importance or salience of issue scale no. k, and a_j^{p} is the party intercept for party j equal to the possible maximum utility of the party, and securing that the predicted individual probabilities add up to the election result.[3]

For multiple issues the proposed alternative model to equation (9) is the directional model

$$u_{ij}^{d} = a_j^{d} + \sum_{k=1}^{K} c_k^{d} X_{ki} Y_{kj} ; \quad i = 1,\ldots,n ; \quad j = 1,\ldots,J , \tag{10}$$

where again the "exponent" d is not an exponent but an indication of the directional model. This time the self-placement X_{ki} is multiplied with the party position Y_{kj} for each issue. It is important to note that the value of the neutral position on each issue scale must be equal to 0. Thus, high utility can be obtained if the self-placement and the party position have high absolute values with the same sign (either positive or negative), but they need not have the same value as in the proximity model to give high utility.

Now, the difference between the two models is not nearly as great as many authors seem to believe (see for example Westholm 1997). On the contrary, as long as both models include party intercepts they both have exactly the same fit to the data. By expanding the squared parentheses in equation (9) we see that

[3] The most common proximity model is a model like this one, which measures the distance between the voter and the party by the squared difference, but other models are also possible. See for example Lewis and King (2000, p. 24).

(11) $$u_{ij}^{p} = \left(a_j^{p} - \sum_{k=1}^{K} c_k^{p} Y_{kj}^{2} \right) - \left(\sum_{k=1}^{K} c_k^{p} X_{ki}^{2} \right) + \left(2 \sum_{k=1}^{K} c_k^{p} X_{ki} Y_{kj} \right) \ .$$

It is crucial to notice that the first parenthesis only varies with *j*, that the second parenthesis only varies with *i* and that the third parenthesis varies with both *i* and *j*. We have already proved above with equation (6) that an additive idiosyncratic factor that only depends on *i* cancels out in the multiple logit model (2). This means that by using the multiple logit model there is no difference between equation (11) and another equation where the second parenthesis is removed from equation (11). Thus, the proximity model (9) can as well be written

(12) $$u_{ij}^{p} = \left(a_j^{p} - \sum_{k=1}^{K} c_k^{p} Y_{kj}^{2} \right) + \left(2 \sum_{k=1}^{K} c_k^{p} X_{ki} Y_{kj} \right).$$

This looks very much like the directional model. Actually, it *is* the directional model if we set

(13) $$a_j^{d} = a_j^{p} - \sum_{k=1}^{K} c_k^{p} Y_{kj}^{2} \ ; \quad j = 1, \ldots, J$$

and

(14) $$c_k^{d} = 2 c_k^{p} \ ; \quad k = 1, \ldots, K \ .$$

This means that the proximity model (9) gives exactly the same fit to the data as the directional model (10) and the one model can be derived from the other. For example, one can either estimate the coefficients in the directional model (10) directly, or first estimate the coefficients in the proximity model (9) and then compute the coefficients in the directional model by using equations (13) and (14). The results will be the same.

In spite of the same fit of the two models this does not mean that the interpretation of the two models is the same. Certainly, there is no interesting difference between the coefficients c_k^{p} and c_k^{d} for the importance of a scale no. *k* in respectively the proximity and the directional model, since the second is just two times the first. More interesting is the difference between the party intercepts in the two models since according to equation (13) the intercept a_j^{d} for party *j* in the directional model can be derived from the intercept a_j^{p} for the same party in the proximity model by subtracting positive values that are larger the more extreme the party is on each scale. This amounts to the "penalty" for extreme parties mentioned above and the penalty d_j for party *j* can be measured by

(15) $$d_j = \sum_{k=1}^{K} c_k^{p} Y_{kj}^{2} \ ; \quad j = 1, \ldots, J \ .$$

At the outset we have no reason to prefer the one model for the other. However, we will prefer that model, which gives the best interpretation of the party intercept as a measure of general popularity of the party.

Denmark 1998. Estimating models for issue voting

Our first example is an analysis of data from the most comprehensive study on issue voting in Denmark concerning the national election to the Danish parliament in 1998. It includes self-placement and voter assessment of party position for ten parties concerning general left-right

position, refugee intake, size of public sector, green politics, treatment of criminals, and further integration with the European Union. The proximity model is the discrete choice model (1), with fixed part from equation (9), and the directional model is the discrete choice model with fixed part from equation (10). Table 1 shows the estimates of the two models using the voter study from the Danish national election in 1998. [4]

Table 1 Proximity model and directional model for party choice at the Danish national election 1998.

	Proximity model			Directional model			Aggregate data	
	Coef.	Std. Err.	Sign.	Coef.	Std. Err.	Sign.	Share	Symp.
Party intercept	a^p			a^d				
United List	-0.975	0.093	0.000	-3.039	0.149	0.000	0.027	-0.260
Socialist P. Party	-0.418	0.067	0.000	-0.904	0.092	0.000	0.076	-0.078
Social Democrats	0.675	0.025	0.000	1.535	0.050	0.000	0.361	0.088
Social Liberals	-1.564	0.026	0.000	-0.593	0.052	0.000	0.039	-0.064
Center Democrats	-1.482	0.019	0.000	-0.501	0.049	0.000	0.043	-0.057
Christian P. Party	-2.150	0.022	0.000	-1.063	0.046	0.000	0.025	-0.156
Conservatives	-0.687	0.029	0.000	-0.778	0.039	0.000	0.090	-0.027
Liberals	0.425	0.032	0.000	0.027	0.045	0.551	0.241	0.043
Danish P. Party	-0.203	0.075	0.007	-2.452	0.139	0.000	0.074	-0.248
Progressive Party	-1.323	0.077	0.000	-3.520	0.134	0.000	0.024	-0.275
Issue saliency	$2\times c^p$			c^d				
Left-right	22.599	1.703	0.000	22.599	1.703	0.000		
Refugees	7.523	0.933	0.000	7.523	0.933	0.000		
Public sector	5.817	0.929	0.000	5.817	0.929	0.000		
Green politics	4.362	1.173	0.000	4.362	1.173	0.000		
Law and order	2.975	0.619	0.000	2.975	0.619	0.000		
European Union	4.883	0.356	0.000	4.883	0.356	0.000		
Pseudo R^2	0.383			0.383				

For the sake of illustration of the similarity of the proximity and the directional model we will in the following estimate the two models separately, although we know that the one can be derived from the other. However, we will multiply the coefficient c_k^p in the proximity model with 2 so that it can be directly compared with the coefficient c_k^d in the directional model.

[4] Notice that the model in this and the following cases is estimated after the issue-scales are transformed to scales from –0.5 to +0.5 to get more readable results with large coefficients. The models and the standard errors are estimated using post-stratification by party choice, so that the individual estimated probabilities adds up exactly to the party shares of the election results for those parties included in the model (two very minor parties were excluded because of lacking information). Post-stratification by party explains the low standard errors of the party intercepts. Each model were estimated with Stata using the last party J as base, but since an arbitrary constant can be added to all intercepts the intercepts were identified by setting the average party intercept weighted with the national shares for each party (shown in the second last column) equal to zero. Thus the base party (Progressive Party) can also be included in the table. The idea is to get more stable intercepts to be used in dynamic analysis. The squared standard error of each transformed intercept is computed from the $(J\text{-}1)\times(J\text{-}1)$ covariance matrix **C** of the original estimated intercepts by the equation **w'Cw** where **w** for the base party J is a column vector of negative party shares for the first J-1 parties. For another party j, the no. j element in **w** is substituted with 1 minus the party share. As an external measure of general popularity the mean sympathy (feeling thermometer) on a scale from -0.5 to +0.5 across all respondents for each party is listed in the last column of table 1.

The same fit of the two models as indicated by the same Pseudo R-square of 0.383 is quite good.[5] As mathematical derived above the issue saliency coefficient c_k^d in the directional model is exactly two times the issue saliency coefficient c_k^p in the proximity model, and the fit of the two models are identical. The tables shows that the saliency of the general left-right scale is much higher than the saliency of politics concerning refugees, the size of the public sector, green politics, law and order, and concerning strengthening the EU, although all the saliency coefficients are significant different from zero.[6]

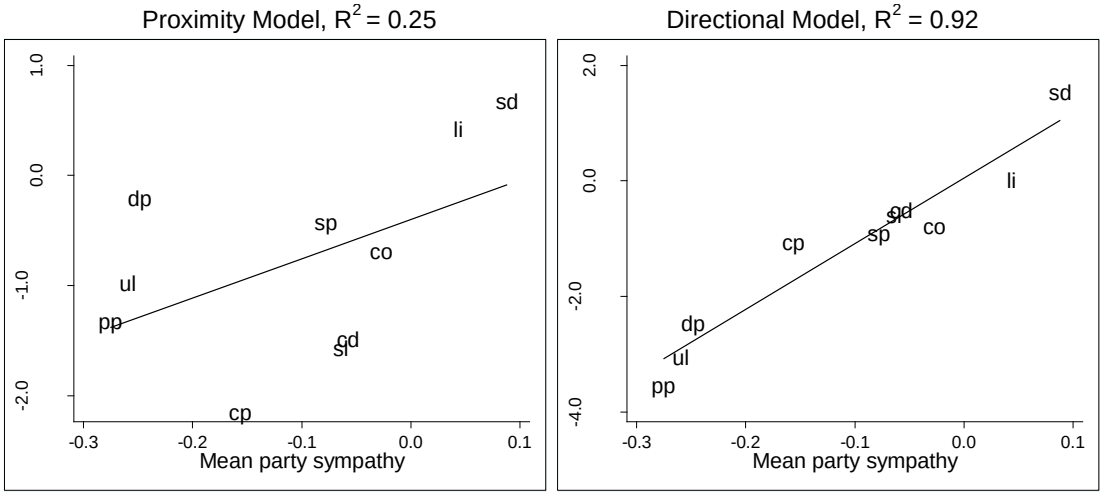
As also mathematical derived the party intercepts are different between the two models, although the party intercept a_j^d in the directional model can be computed from the party intercept a_j^p in the proximity model using equation (13). Since an arbitrary constant can be added to all intercepts they are all measured as deviations from a common average (see footnote 4). Because of the penalty in the directional model, the extreme parties, such as the left-wing United List and the right-wing Danish People's Party and Progressive Party, have more negative party intercepts in the directional model than in the proximity model. As mentioned above, the different party intercepts in the two models are alternative measures of general popularity. In figure 1 the two sets of party intercepts are compared with an external measure for general popularity i.e., the mean sympathy for each party on a "feeling thermometer" scale.

In the Danish example the relation between party intercept and mean party sympathy is much stronger linear for the directional model than for the proximity model. In the graph to the left there is no clear pattern, and the size of the intercepts mostly indicate the size of the party shares. This also indicates that the spatial proximity model cannot sufficiently account for the size of the parties without the party intercepts. In contrast, the graph to the right for the directional model shows that the extreme left-wing party (ul) and the extreme right-wing parties (pp and dp) to a certain extent can compensate for low sympathy by being extreme.

[5] We use the McFadden Pseudo R^2 that is $1-L_1/L_0$ where L_1 is the log likelihood of the full model while L_0 is the likelihood of a model where all respondents have the same probability for voting for all the parties. Thus, the closeness of pseudo R^2 to 1 indicates how close we are to perfect prediction of the choices of all the individual voters where the log likelihood is equal to 0 (McFadden 1974). This fit measure is quite conservative compared with other types of pseudo R^2.

[6] It is probably a problem that the left-right scale is a kind of meta-scale encompassing all the other scales and it is thus taking explanatory power from the other scales. It is included anyway to get a good overall measure of how extreme each party is in case important issues are accidentally ignored. The strategy in the following is to always include all issues with self-placement and voter-assessment of party positions in a certain study, unless they are insignificant or with opposite sign than expected.

Figure 1 Denmark 1998. Party intercept by mean sympathy in both the proximity and directional model



cd: Center Democrats; cn: Conservatives; cp: Christian People's Party; dp: Danish People's Party; lb: Liberals; sd: Social Democrats; sl: Social Liberals; sp: Socialist People's Party; pp: Progressive Party; ul: United List.

To see if these results just occurred by coincidence we will in the following estimate the two models on other election studies with good data about issue voting i.e., data about self-placement and assessment of party positions on several issues, and about party sympathy.[7]

Two other proportional systems: Norway 1989 and Sweden 1994

In this section we will study issue voting in two other Nordic countries with multiparty systems and strong proportional electoral systems: Norway and Sweden.

The Norwegian study of the parliamentary election of 1989 is probably the most analysed study in the literature on issue voting including data on general left-right position, farming subsidies, green politics, alcohol policy, immigration, health and crime for seven parties, excluding some minor parties. With a pseudo R-square of 0.40 the fit of both individual-level models is as good as in the Danish case.

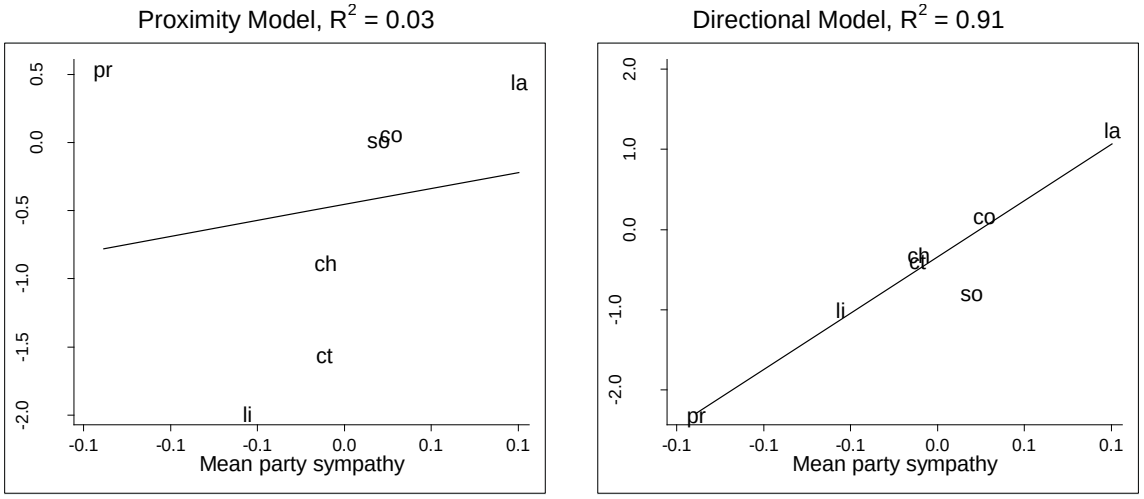
Figure 2 also shows a similar pattern as in the Danish case with a strong linear relation between party intercept and mean party sympathy for the directional model, and an almost absent relation for the proximity model. To my knowledge no one has before demonstrated this clear relationship on the aggregate level either for the Norwegian data or for any other country. In Norway it is especially the right-wing Progress Party that can compensate for low sympathy by being extreme on the immigration policy issue.

The Swedish data on issue voting at the parliamentary election of 1994 is from the Swedish Election Study 1994. It includes data on issue voting about the general left-right dimension, green politics, EU, refugees and gender equality. In this case the pattern of relationships in Figure 3 between party intercepts and mean party sympathy is not as clear as in Denmark and Norway. However, the linear relationship is still closer for the directional model than for the proximity model and it appears that especially the right-wing party New Democracy (nd) could compensate for low sympathy by being extreme. It is especially the quite high sympathy with the Green Party that is not reflected very well in the support by the voters indicated by the party intercept. One explanation could be that green policy is becoming an uncontroversial valence issue as "something good", but without necessarily qualifying for electoral support. But this put the concept of "general popularity" under scrutiny. What is it actu-

[7] Unfortunately, data from the the Comparative Study of Electoral Systems (CSES) are not appropriate since they only include a single general issue dimension for each country.

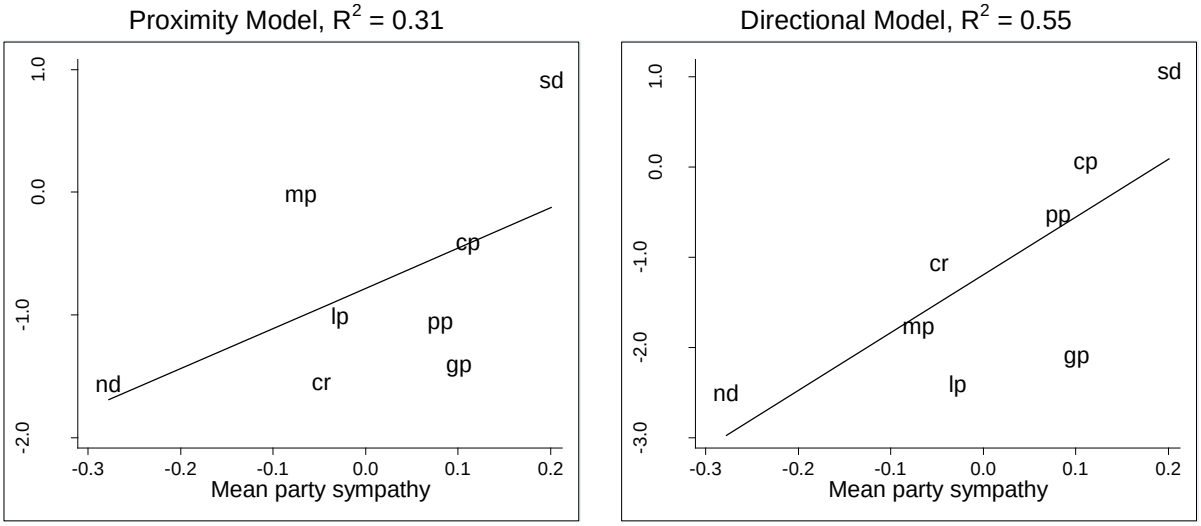
ally besides position politics that gives a general intensive to vote for a certain party? It is obviously not always mere sympathy.

Figure 2 Norway 1989. Party intercept by mean sympathy in both the proximity and directional model



so: Socialist Party; la: Labour; li: Liberal Party; ct: Centre Party; ch: Christian People's Party; co: Conservatives; pr: Progress Party

Figure 3 Sweden 1994. Party intercept by mean sympathy in both the proximity and directional model



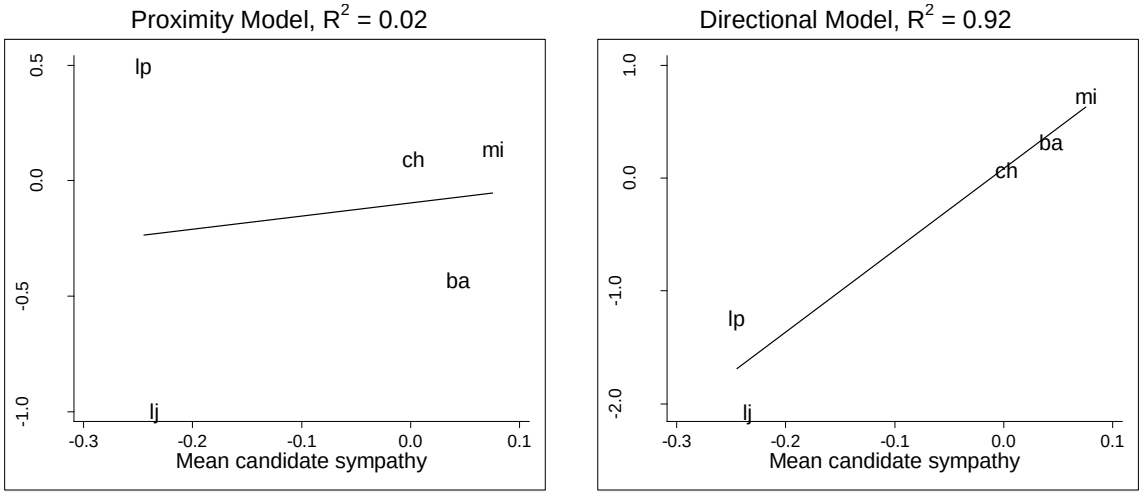
lp: Left Party; sd: Social Democrats; gp: Green Party; pp: People's Party; cp: Centre Party; cr: Christian Democrats; mp: Moderate Party; nd: New Democracy

Other similar to proportional systems: France and Ireland.

France do not has a proportional electoral system as the Nordic countries, but the first round in the two rounds for the election of the president bear some similarity to the Nordic proportional systems at the national elections, because the voters can vote for the same set of multiple candidates in every district. Issue voting is not included in the French electoral studies by the same approach as described above, but a single study using this approach was made by an American scholar (Pierce 1997) at the 1988 presidential election. It included issue voting concerning the general left-right dimension, subsidies for church schools, size of the public sector

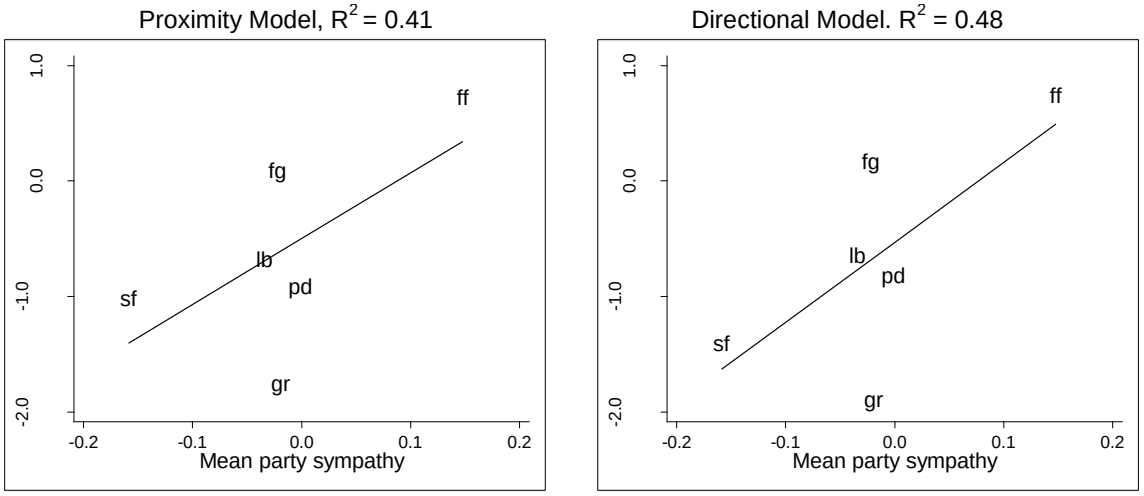
and immigration policy for five major candidates. Interestingly, figure 4 shows that the pattern of relationship between party intercepts and mean party sympathy with only five data points is as clear as in Denmark and Norway. It is especially the left-wing communist candidate Lajoinie and the right-wing Le Pen who compensated for low mean sympathy by being extreme.

Figure 4 France 1988. Party intercept by mean sympathy in both the proximity and directional model



lj: Lajoinie mi: Mitterand ba: Barre ch: Chirac lp: LePen

Figure 5 Ireland 2002. Party intercept by mean party sympathy in both the proximity and directional model



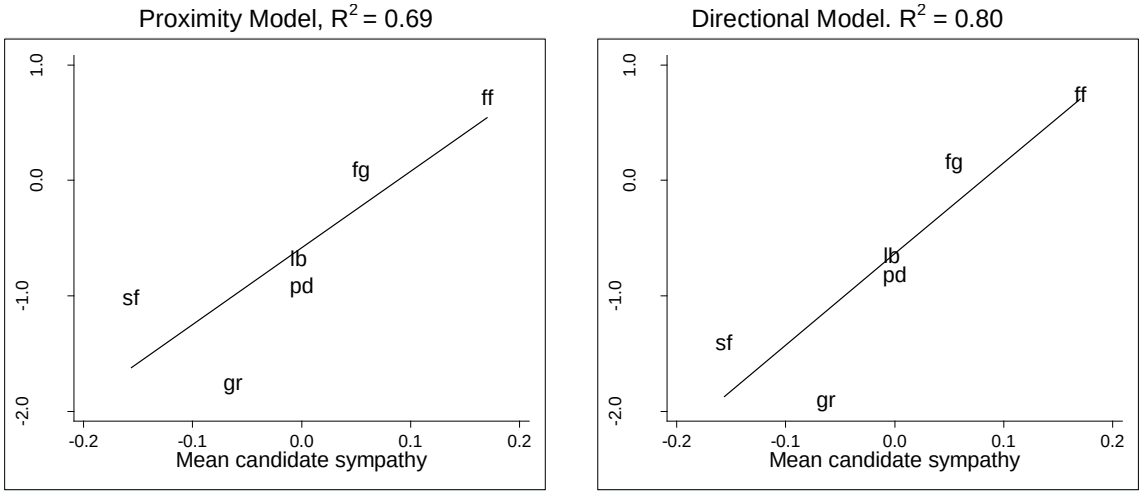
ff: Fianna Fail; fg: Fine Gael; gr: Green Party; lb: Labour; pd: Prog. Dem. sf: Sinn Fein

The electoral system in Ireland is proportional representation by Single Transferable Vote that combines a fair representation of the partisan distribution at the constituency level with the opportunity of ranging the local candidates after preference. The data analysed is the Irish Election Survey for parliamentary election of 2002 (Marsh et al., 2008). It was the first comprehensive election survey in Ireland including questions about issue voting, party identification and sympathy with parties, leaders and candidates. Especially the sympathies with local candidates were carefully investigated by asking about the sympathy with named major candidates in the respondent's constituency.

The following analysis concerns first preference vote indicating the party of the first chosen candidate. The six major parties were included in the study. The Irish party system is based on post-independence political factions, and politics are still much influenced by inter- and intra-party factional and personalised strife instead of ideological battles and policy disputes. Thus the traditional left-right scale has no special meaning in Ireland and was not included among the issues. The many votes for independent candidates (about 10 pct. of the votes) were excluded from the analysis. The included issues were about environment protection, EU integration, United Ireland, abortion and government spending.

As expected from the low importance of policy-voting in Ireland the explanatory power of issue-voting was not very high (pseudo R-square = 0.18), and the relation between party intercepts and mean party sympathy in Figure 5 is not very strong, although respectable. Interestingly, the graph for the proximity model in the left panel of Figure 5 is quite similar to the graph in the right panel for the directional mode. The explanation is that the Irish parties are not very extreme on the major issues and thus the directional model tends to converge towards the proximity model[8]. The graph for the directional model is strangely reminding of the Swedish results in figure 3 since the sympathy for the Green Party is not very much reflected in the electoral support. As an opposite case the support for the Fine Gael is quite high in spite of the modest sympathy for the party at the national level. This raises the question if the sympathy with the local candidates can better explain the support for the parties. This seems actually to be the case. Because the Irish election study has data on the sympathy with local candidates we can compute sympathy with the most favoured local candidate from each party as an alternative measure to party sympathy. With this measure the relation between party intercepts and mean sympathy with the candidates of each party in Figure 6 becomes nearly as strong for the directional model as in the French case. Because of the moderation of the Irish parties the graph for the proximity model does not differ much from the graph for the directional model. However, Sinn Fein that is the political wing of the IRA, compensated for the general low sympathy with the Sinn Fein candidates by being extreme on the United Ireland issue.

Figure 6 Ireland 2002. Party intercept by mean candidate sympathy in both the proximity and directional model



ff: Fianna Fail; fg: Fine Gael; gr: Green Party; lb: Labour; pd: Prog. Dem. sf: Sinn Fein

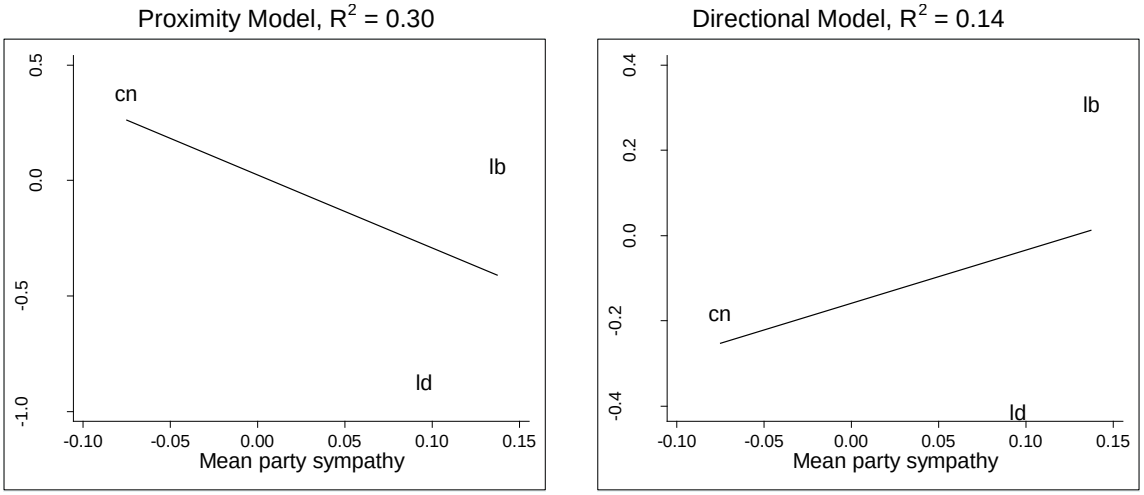
[8] This can be inferred from equation (13). When the party positions Y_{kj} are not very extreme (i.e., not very different from 0) then the intercept in the directional model converges towards the intercept in the proximity model.

Majoritarian systems: UK and the USA

We finally come to the most difficult cases for using multi-party individual level models: the majoritarian system where the winner takes all in the local constituencies (states in US presidential elections) and thus tend to shrink to two party systems (Duverger 1954). With two data points one can always find linear relations, but an election with a quite strong third party in the UK and a quite strong third candidate in the USA were selected.

In UK it was the 1997 General Election where the Liberal Democrats had quite strong following compared with the two major parties Labour and Conservatives. Because of special regional parties in UK outside England, only data from England was selected. The study included issue-data on the general left-right dimension, taxes versus spending, income equality, price control and privatisation of nationalised industries. Not included as insignificant were the issue of gender equality. The relation with only three data points between party intercepts and mean party sympathy in Figure 7 was scattered and negative for the proximity model, but at least positive but very scattered for the directional model. Actually, the results with the directional model make good sense considering the electoral system. It is well known that even quite popular third candidates have difficulties being elected in majoritarian systems and thus the results of the directional model clearly shows the difficulty of transforming general sympathy to general support. Thus the Liberal Democrats only got 19 pct. of the votes among the three parties in spite of a mean sympathy close to that of Labour an much larger than the sympathy with the Conservatives. In this way the directional model might be an instrument for measuring the difficulties in majoritarian systems with transforming sympathy to electoral support.

Figure 7 England 2002. Party intercept by mean sympathy in both the proximity and directional model



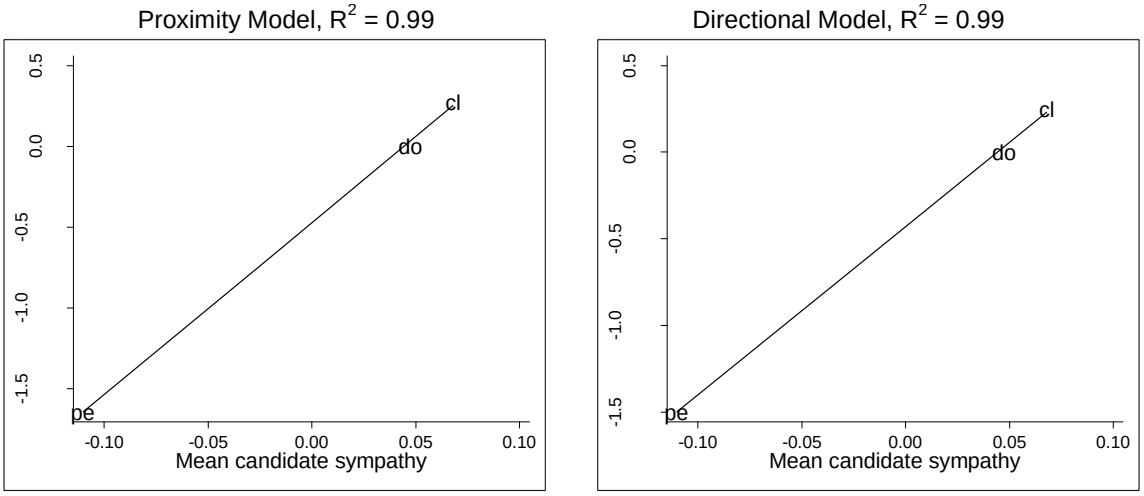
lb: Labour; ld: Liberal Democrats; cn: Conservatives

The United States have a few presidential elections with relative strong support for third candidates, among them the 1996 presidential election where Ross Perot got 8.5 pct. of the votes.[9] The National Election Study for 1996 included a lot of issue data concerning the liberal-conservative scale, role of government, defence spending, private health service, job guarantee, abortion and environment regulation. The following issues were removed as being insignificant in the individual-level model: help for blacks, being tough on crime, jobs versus

[9] I would have preferred to analyse the 1992 presidential election where Ross Perot got 19 pct., but unfortunately the National Election Study for this election only included the single liberal-conservative scale.

environment regulation and equal rights by gender. The results in Figure 8 are actually a bit surprising, since I would have expected a picture similar to the English case in Figure 7. Both the graph for the proximity model and the graph for the directional model show a perfect linear relation, which is not surprising considering that both the sympathy and the intercept is very close for Dole and Clinton. The surprising part is that the slope of the line is as high as for the Nordic countries indicating the same effect of sympathy on general support. Nearly the same pattern is found in the graph to the left for the proximity model and in the graph to the right for the directional model indicating that the positions of the candidates are extremely moderate on all issue scales which is in strong contrast to Northern Europe. On many issue scales the position of Perot was somewhere between Dole and Clinton. The main lesson from this analysis might be that the US majoritarian system does not destroy the relation between sympathy and support, but it creates policies of all candidates closer to the centre, as predicted by Downs (1957)

Figure 8 USA 1996. Party intercept by mean sympathy in both the proximity and directional model



cl: Clinton; do: Dole; pe: Perot

Conclusion

The conclusion from the analysis so far is that the directional model is in fact superior to the proximity model in separating the general support created by valence issues from the individual support created by position issues. This is suggested by the close linear relation between mean party sympathy and the party intercepts estimated by the directional model. In fact it is surprising how close this relation is in many cases considering that the results by an otherwise valid model might by tarnished by at least three possible problems:

- Selection bias. All relevant issues might not be included in the election study.
- Midpoint bias. The neutral point on an issue scale might no be the real status quo point in actual politics. One approach to solve this problem might be to use the mean self-placement of all voters as the neutral point instead of the midpoint on the scale.
- Time bias. The election survey is not done exactly on Election Day and might not be finished before several weeks after the election. In this situation the attitudes of the voters, the sympathy with the parties, and the preferred party might not be the same as on Election Day.

The strong results in this paper, in spite of these possible problems, calls for further studies on the effect of valence issues, preferable by analysing both several consecutive elections within the same country and by analysing several countries with different electoral systems.

The main lesson is that voting behaviour cannot be understood without considering the common interests among all voters together with the more special interest of the individual voters. Thus, from the voters' point of view politics is not only influenced by positions in a policy space, but also by the shared values that define the culture of a society.

There are strong indications that the directional model for voting behaviour is better in separating the effect of common interest concerning valence issues from the effect of special interests concerning positional issues. It is especially in the European multi-party systems with proportional representation that the general sympathy with the different parties are reflected in the election results, while this is less clear in majoritarian systems that forces the voters to act on other considerations than sympathy.

References

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

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

Butler, D.E., Stokes, D., 1981. Political Change in Britain: Forces Shaping Electoral Choice. Comparative Political Studies 14 (1), 31-47.

Butler, D., Van Beek, S.D., 1990. Why Not Swing? Measuring Electoral Change. PS: Political Science and Politics , 178-184.

Dow, J.K., Endersby, J.W., 2004. Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research. Electoral Studies 23 (1), 107-122.

Downs, A., 1957. An Economic Theory of Political Action in a Democracy. Journal of Political Economy 65 (2), 135-150.

Duverger, M., 1954. Political Parties, their Organization and Activity in the Modern State.

Inglehart, R., 1997. Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies.

Karandikar, R.L., Payne, C., Yadav, Y., 2002. Predicting the 1998 Indian Parliamentary Election. Electoral Studies 21 (1), 69-89.

Lewis, J., King, G., 2000. No Evidence on Directional Vs. Proximity Voting. Political Analysis 8 (1), 21-33.

Long, J.S., 1999. Regression Models for Categorical and Limited Dependent Variables. Social Forces 77 (3), 1245-1246.

Long, J.S., Freese, J., 2006. Regression Models for Categorical Dependent Variables using Stata. 2. ed.

Marsh, M., Sinnott, R., Garry, J., Kennedy, F., 2008. The Irish Voter: The Nature of Electoral Competition in the Republic of Ireland.

Merrill,Samuel, I.I.I., Grofman, B., 1999. A Unified Theory of Voting: Directional and Proximity Spatial Models. , 213.

Pierce, R., 1997. Directional Versus Proximity Models: Verisimilitude as the Criterion. Journal of Theoretical Politics, Vol.9, Issue 1, p.61 9 (1), 61.

Rabinowitz, G., Macdonald, S.E., 1989. A Directional Theory of Issue Voting. American Political Science Review 83 (1), 77-91.

Schofield, N., Sened, I., 2006. Multiparty Democracy: Elections and Legislative Politics. , 223.

Westholm, A., 1997. Distance Versus Direction: The Illusory Defeat of the Proximity Theory of Electoral Choice. American Political Science Review 91 (4), 865-883.