

¹A Comparison of Different Estimation Methods of Voting Transitions with an Application in the Dutch National Elections of 2003 and 2006

Carin van der Ploeg¹, Frank van de Pol², Jarl Kampen³

¹Statistics Netherlands, e-mail: ce.vanderploeg@cbs.nl

²Statistics Netherlands, e-mail: f.vandepol@cbs.nl

³Wageningen University, e-mail: jarl.kampen@wur.nl

Abstract The estimation of changes in voting behaviour can be pursued in two ways, by modelling aggregated election results, and by means of recall data recorded in survey questionnaires. In an application of these methods for the Dutch national elections of 2003 and 2006, we show that the estimated voting transition estimated by survey techniques and model based techniques complement each other, improve the validity of the results, and provide a basis for new research.

Keywords: Voting transitions, transition matrix, maximum likelihood

1. Introduction

Just like politicians and the media, political scientists are very keen on the transitions in party preference of voters between elections. This is visible in the large amount of literature on *movers* and *stayers* that can be found in the literature and goes back at least to Leo Goodman's article from 1961, *Statistical methods for the mover-stayer model*. Also, many Dutch opinion polls varying from Maurice de Hond, Interview NSS to TNS NIPO periodically publish predictions of political transitions and support for political parties. The goal has always been to estimate the voting behaviour as precisely as possible. This is a big challenge due to the confidentiality of the voting ballot.

The goal of the present study is to improve and compare methodology regarding the estimation of voting transitions. A comparison is made between model based methodology and survey based methodology. The election results from two consecutive elections are used to estimate voting transitions between two subsequent elections. The first occasion is the election year of 2003 and the second occasion is the election year of 2006. The quality of the estimated transition matrices is evaluated by comparing estimated voting results at the second occasion that are implied by the first election results and the transition matrix to be evaluated. These estimates for the second occasion can be compared to the observed election results. We note in passing that this methodology of estimating election results by means of a transition matrix is applicable in a broader field of cohort studies with categorical variables.

¹ The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands.

The remainder of this paper is structured as follows. In section 2 the state of the art of the established methodology is presented. In the third section the results are presented, and the final section provides a discussion and topics for further research.

2. Methodology for describing mover-stayer behavior

2.1 Basic notation

Consider two subsequent elections, with n parties coming up in the first election and m parties coming up at the second election. The observational units are defined by municipalities, denoted by index g . The following basic formula is the foundation for this paper:

$$\Pr(j | g) = \sum_i^n \Pr(i | g) \times \Pr(j | i), \quad (1)$$

where the parties in the first election are represented with index $i=(1, \dots, n)$, and the parties in the second election with index $j=(1, \dots, m)$. Further, $\Pr(i | g)$ denotes the observed proportion of voters that voted for party i in municipality g at the first occasion, and $\Pr(j | g)$ is the observed proportion that voted party j at the second occasion, in given municipality g . These proportions are given. However, the actual proportion of voters that move from party i to party j , denoted, $\Pr(j | i)$, are *unobserved* parameters. From the formula

$$\Pr(j | i) = \Pr(i, j) / \Pr(i), \quad (2)$$

the challenge is to estimate the underlying joint probabilities $\Pr(i, j)$. These cannot be observed directly from aggregated data and have to be estimated (see sections below). Note that in line with e.g., Keller & ten Cate (1977), a single (national) transition matrix is assumed to hold for all municipalities,

$$\Pr(j | i, g) = \Pr(j | i). \quad (3)$$

2.2 Benchmark methods

With $\Pr(i, j)$ unknown, for starters we may simply fill in values, and evaluate the resulting estimates of $\Pr(j | g)$ by applying these transition probabilities in Formula (1) and comparing these estimates with the observed values. For instance, we may assume the independence model or perfect mobility model, meaning that the votes of a constituent at the second occasion is totally independent from its vote at the first occasion. This can be represented by this formula (for an introduction to the mover-stayer model, see Goodman, 1961):

$$\Pr(j | i) = \Pr(j). \quad (4)$$

Alternatively, the 100% stayers model may be considered, meaning that voters are assumed to be perfectly loyal to the party at the first election:

$$\Pr(j | i) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}. \quad (5)$$

These extreme models are useful for comparison and evaluation of advanced approaches for estimating transitions, which are discussed below.

2.3 The survey approach

The Dutch Election Survey (Nederlands Kiezers Onderzoek; NKO) is held around every national election. The last survey has been carried out by Statistics Netherlands (Centraal Bureau voor de Statistiek). In the NKO respondents are asked what their most recent vote was as well as their vote at the previous election. Using these data, expressing the transitions between parties back and forth proceeds in a straightforward manner. For the 2003-2006 elections, the established transition matrix can be found in Aarts *et al.* (2007: pp. 224) and is reproduced here in Table 1, where like Aarts *et al.* (2007) the sample records have been multiplied by a weight factor in order to obtain the most truthful representation of the Dutch Population².

As said, the thus obtained transition matrix can be used to predict electoral results in 2006 from electoral results in 2003, by applying Formula (1). However, it is well established in the literature that recall data can be very unreliable (Weir, 1975). There are several effects making accounts of respondents unreliable (Voogt, 2004; Keller & ten Cate, 1977; Upton, 1977).

- Respondents do not recall their previous vote.
- Respondents do not want to be fickle, so there is a bias towards answering the same choice of party between the previous and current election.
- Respondents answer incorrectly and for instance name the party for which they voted in the municipality, provincial or European elections.
- Respondents say that they have voted but in fact didn't vote in the previous election, wanting to give a sociable accepted answer.
- Respondents want to belong to the winning side and answer incorrectly to have voted a party who won in the elections.

It is not easy to determine exactly what the magnitude of this bias is. According to Voogt (2004), the most important effects are non-response bias, response bias and stimulus effect. Non-response bias is the most important effect and accounts for an underrepresentation of certain population groups. Extensive literature on this problem shows that in general non-respondents are higher educated, younger, more often single with an overrepresentation in the urban areas (see for a list of published literature Voogt, 2004: pp. 35). The response bias, also called answer and memory effects, reflects the above mentioned reasons for unreliable results (Bethlehem & Kersten, 1986). Such bias occurs especially in panel research, where respondents are interviewed at least twice (Greenwald *et al.*, 1987). People that did not intend to cast their vote, nevertheless tend to vote when they have participated in a pre-election poll. For all of these reasons, alternative strategies for estimating transition matrices are very much desirable.

2.4 Maximum Likelihood Estimation

Several algorithms based on least squares regression can be used to estimate transition probabilities, as was shown by Keller and Ten Cate (1977) and Van der Ploeg (2008). Avoiding negative estimates of proportions takes special precautions in that approach. An alternative methodology produces maximum likelihood (ML) estimates of the

² The values of this transition matrix differ slightly from the values in the book of Aarts *et al.* (2007) because the matrix in this paper has values that are rounded such that row totals add up to 1. This is not the case with the matrix presented in the book of Aarts *et al.*

transitions based on the observed marginal probabilities $\Pr(i | g)$ and $\Pr(j | g)$. We prefer the latter because it optimises the likelihood and therefore produces better fitting estimates of the second election. Elaborating on basic theory supplied by Clogg & Goodman (1984) and Van de Pol and De Leeuw (1986), maximum likelihood estimation iteratively produces the most likely values of the conditional probabilities $\hat{\Pr}(j | i)$, given the election results on both occasions. Assuming the complete table $\Pr(i, j, g)$ to be known, the likelihood function (L) under the assumption that $\hat{\Pr}(j | i)$ is given by

$$L = \prod_{i,j,g} \hat{\Pr}(i, j, g)^{N\Pr(i,j,g)} = \prod_{i,j,g} \Pr(g)^{N\Pr(g)} \times \Pr(i | g)^{N\Pr(i|g)} \times \hat{\Pr}(j | i)^{N\Pr(j|i,g)}, \quad (6)$$

with N the total size of the voting registry, averaged over both occasions. Because of the analytical infeasibility of direct calculation of the maximum likelihood function, we apply the Expectation Maximization (EM) algorithm (Dempster, Laird & Rubin, 1977). Firstly, the conditional expectation of the complete-data log-likelihood function (formula 6) is given, based on the observed data and an initial expectation of the missing parameters. Second, these parameters are updated, summing over municipalities, so that the expectation is maximized. Dempster *et al.* (1977) already pointed out that this algorithm in some cases converges to local optima. In order to find the global maximum of the likelihood, we therefore use a large number of sets of starting values.

3. Results

3.1 Transition matrices

Table 1 shows where, according to the NKO survey based on recall data, the voters of 2003 have gone in 2006. For example, 71% of the voters who voted for the Christian-Democrats (CDA) in 2003 also did so in 2006, whereas 3% moved to Labour Party (PvdA) and 6% to the Liberal Party (VVD) etc. Also, 70% of the Socialist Party (SP) voters of 2003 stayed with the SP in 2006, but 11% voted for the PvdA instead. On the other hand 22% of the PvdA voters of 2003 voted for the SP in 2006. The relatively high values on the diagonal mean that most voters were loyal to their party. They are called stayers. Democrats 66 (D'66) has lost a lot of support and there are relatively few stayers in this party in comparison to other parties. It becomes also quite visible that a lot of former voters for the List Pim Fortuyn (LPF) a former populist right wing party have voted in 2006 for the Party For Freedom (PVV), a new right wing party.

Table 2 contains the transition probabilities as estimated by maximum likelihood (see Section 2.4). Inspection of the ML transition matrix reveals that most voters are stayers. Especially the SP, CDA, PvdA, Christian Union (CU), VVD and Green Left (GL) have high stayer rates. Someone who did not vote in 2003 in most cases also abstained in 2006 (81%). The proportion of non-voters (NV) is a disputable piece of information, even with the model based techniques. You can still see that 11% of the non-voters in 2003 did vote for the SP in 2006, explaining partly the growth of the SP. The 2003 electorate of D'66 was fragmented over at least four parties. According to the ML model, in 2006 over 55% of the previous LPF voters casted their vote in favour of the PVV and a large portion of former LPF voters remained at home. The SP and CU have values of 1 on their diagonal indicating that they did have very loyal support.

3.2 Goodness of fit and Pseudo-R² statistics

To summarize discrepancies between predictions and actual election results, we apply the well known log-likelihood ratio statistic

$$LRX^2 = 2N \sum_g \sum_j \Pr(j | g) \times \ln \left(\frac{\hat{\Pr}(j | g)}{\Pr(j | g)} \right) . \quad (7)$$

McFadden's pseudo R² (McFadden, 1973, pp. 121) is used to show the relative fit of each of the estimated transition matrices (100% stayer, NKO survey, and maximum likelihood) as compared to the independence (perfect mobility) model. Larger values of the R² indicate a better fit of the involved model. When R²=1, with the given transition matrix, the second election results at the municipality level can be predicted perfectly from the first election results.

Table 3 shows a significant improvement of the fit by the NKO matrix in comparison to the independency model, and in comparison to the 100%-stayer model. The NKO pseudo-R² equals 0.920, or 92% of the deviance of the independence model is explained by the corresponding transition matrix. The ML-model however, performs even better, with a pseudo-R² of 0.967. In Figures 1a-c, we provide a visual presentation of the pseudo-R² statistics per municipality. Lighter colouring means higher values and thus a better fit. The divisions in colour are equal to the 20th, 40th, 60th and 80th percentile of the pseudo-R² statistics derived from the NKO. Only in a few municipalities the 100%-stayer model is able to adequately predict the voting behaviour of 2006 from the voting behaviour in 2003. To be precise, in Staphorst, Urk en Bunschoten a good prediction can be found (McF. pseudo R²>.98). The so called 'bible belt' is very clearly visible and accounts for very loyal voters in these municipalities (McF. pseudo R²>.81). The map of The Netherlands becomes considerably lighter if the NKO-matrix is used, especially for the provinces of Northern Brabant and Overijssel. But in the most urbanised parts of the Netherlands and Northern and Southern Holland the results are quite poor. The ML-matrix is the best model to predict the shifts. This is to be expected because in contrast to the survey, ML is a process of optimisation.

Finally, all approaches have problems to predict the voting transitions for the islands correctly. The explanation for this lies within the fact that the islands are popular resorts for holidays. The population of the islands changes considerably depending on the tourists that are on holiday there. Also all models do not estimate Limburg correctly. Further analysis shows that the support for the PVV in Limburg is systematically underestimated. Taking into account that the founder of the PVV is from this part of The Netherlands, one can easily understand why Limburg deviates from the national pattern.

4. Discussion and further research

The most important assumption of the benchmark model of independence is that there is perfect mobility of a voter between two elections. In this model, the probability to vote for a certain party on the second occasion is independent of the choice of party on the first occasion. Analysis shows that with this model the election results on the level of municipality cannot be predicted using this model. As expected, also the model that assumes that all voters are completely loyal does not reflect reality. Using a transition matrix that is estimated from NKO data, improves the model fit considerably. Model based estimation techniques improve these results even further. The results show that on

the level of voting transitions NKO and the model based estimation techniques are very similar.

A great advantage of NKO is that it can map the reported mobility of voters. For each respondent it can be shown what party (s)he voted. Therefore, transitions between different parties can be analyzed in both directions. Model based estimation techniques produce gross transitions based on net transitions between parties. This way, transitions between two parties often diminish, and consequently the transition matrix contains many zeros. This is closely related to the problem of ecological inference (King, 1997; King *et.al.*, 2004).

An advantage of model based research is that it is a cheap technique compared to survey research; and municipal election results are always available. A second advantage is that with model based estimation techniques the total population of the Netherlands can be observed, while the NKO research is a sample survey in which the voting behaviour is extrapolated to the total population. In this way, the model based research made it possible to observe interesting regional behaviour. In the estimations it becomes visible what the influence of certain politicians is on the voting behaviour in the regions where they come from. Also, the difficulty in explaining transitions on the Dutch islands becomes visible. Other purely election specific elements are better visualized. For instance, in 2006 the voting transitions of Limburg clearly deviated from other provinces.

The most important conclusion of this research is that the different methods (sample survey and model based) are in a way complementary. Results of the methods can be confirmed by each other, improving the validity of the conclusions. Furthermore, this research shows that the assumption that there is a single transition matrix that is the basis of the voting behaviour of all Dutch voters is not feasible. Future research will have to be done to investigate the influence of fine tuning of the models for obtaining the transition matrices. Possible directions of research are: firstly, using different matrices for different regions or for different degrees of urbanization. Secondly, other options such as Self Organizing Maps and Stochastic global search methods, may prove to be valuable optimization techniques. Self Organizing Maps is a technique within artificial neural networks. This technique can be used to make logical groups of municipalities which can be used to make a better division in regions and to make multiple transition matrices. Thirdly, it could be interesting to define clusters on the basis of electoral support for a certain influential party. This technique may be more applicable for 2 or 3 party systems, but it would also be possible to be applied in multi-party systems.

Bibliography

- Aarts, K., H. van der Kolk & M. Rosema (2007). *Een verdeeld Electoraat De Tweede Kamerverkiezingen van 2006*, Spectrum.
- Agresti, A. (1996). *An introduction to categorical data analysis*, John Wiley & Sons, New York.
- Bethlehem J.G. & H.M.P. Kersten (1986). *Werken met Nonresponse (Working with Nonresponse)*. Voorburg: Centraal Bureau voor de Statistiek.
- Clogg, C. C. & L. A. Goodman (1984). Latent structure analysis of a set of multidimensional contingency tables, *Journal of the American Statistical Association* **79**:762-771.

- Dempster, A. P., N.M. Laird & D.B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society, Series B, Methodological*, **39**: 1 - 37.
- Goodman, L. A. (1961). Statistical methods for the mover-stayer model, *Journal of the American Statistical Association* **56(296)**: 841-868.
- Greenwald, A., C., Carnot, R. Beach & B. Young (1987). Increasing voting behavior by asking people if they expect to vote, *Journal of Applied Psychology*, **72**: 315-318.
- Keller, W.J., & A. ten Cate (1977). De verschuiving van de kiezersvoorkeur, *Economische Statistische Berichten* 26-10-1977.
- King, G. (1997). *A Solution to the Ecological Inference Problem. Reconstructing Individual Behavior from Aggregate Data*, Cambridge MA, Cambridge University Press.
- King, G., O. Rosene en M. Tanner (2004). *Ecological Inference. New Methodological Strategies*, Cambridge University Press, Cambridge.
- McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior, in Zarembka, P. (ed.), *Frontiers in Econometrics*, pp.105-142, Academic Press, New York.
- Pol, van de F. & J. de Leeuw (1986). A latent Markov model to correct for measurement error, *Sociological Methods & Research* **15**, pp. 118-141.
- Ploeg, C.E. van der (2008). *A comparison of Different Estimation Methods of Voting Transitions with an Application in the Dutch National Elections*. (Vrije Universiteit, Amsterdam)
- Upton, G.J.G. (1977). A Memory Model for Voting Transitions in British Elections, *Journal of the Royal Statistical Society Series A* **140**: 86-94.
- Voogt, R. (2004). *I'm not interested- Nonresponsive bias, response bias and stimulus effects in election research*, Universiteit van Amsterdam.
- Weir, B.T. (1975). The Distortion of Voter Recall, *American Journal of Political Science* **14(1)**: 53-62.

Table 1. NKO transition matrix 2003-2006 (diagonal elements bold) (N=2528)

2003	2006									
	CDA	PVDA	VVD	SP	GL	D66	CU	PVV	Other	NV
CDA	.707	.034	.058	.064	.004	.004	.040	.020	.010	.056
PVDA	.029	.591	.016	.204	.027	.008	.008	.012	.014	.092
VVD	.230	.028	.552	.035	.003	.010	.006	.047	.019	.069
SP	.0378	.106		.697	.061		.015	.038	.023	.023
GL	.032	.074	.011	.253	.463	.011	.042		.021	.095
D66	.081	.174	.174	.151	.116	.233	.023		.047	
CU	.022	.022		.022			.911	.022		
LPF	.071	.035	.177	.141		.012		.365	.047	.151
Other	.095	.024	.024	.095	.024		.071	.024	.571	.071
NV	.060	.062	.027	.087	.004	.004	.002	.047	.011	.696

Source: (Aarts *et al.*, 2007: pp. 224), with slight adaptations to avoid rounding errors.

Table 2. Transition matrix ML model 2003-2006 (diagonal elements bold)

2003	2006									
	CDA	PVDA	VVD	SP	GL	D66	CU	PVV	Other	NV
CDA	.830		.003	.066			.009	.040		.051
PVDA		.742		.203			.015			.040
VVD	.168		.789				.023			.021
SP				1.000						
GL					.839	.161				
D66		.170	.145		.075	.280	.008		.198	.125
CU							1.000			
LPF		.005				.002		.559	.222	.213
Other							.145		.855	
NV		.016		.115				.065		.805

Table 3. Goodness of fit statistics for four models 2003-2006

Model	LRX^2	$DLRX^2$	Ddf	P	McFadden Pseudo- R^2
Independency	3290676		18		
100% Stayers	1956575	1334101	-18	<.0001	0.419
NKO	342285	2948391	72	<.0001	0.920
ML	139716	3150960	72	<.0001	0.968

