Forsvaret finder sted den 8. maj 1987 kl. 13.00 (prc.)
i Auditorium A1 (bygning 333) på Aarhus Universitet

Danish Elections 1920-79

Danish Elections 1920-79

A Logit Approach to Ecological Analysis and Inference

Søren Risbjerg Thomsen

Politica

Denne afhandling er af Fagrådet for Statskundskab antaget til offentligt at forsvares for doktorgraden i statskundskab.

Århus, 19. februar 1987

Palle Svensson
Fagrådets næstformand

ISBN 87-7335-067-2
1. udgave, 1. oplag
Omslag: Søren Kjær
Tryk: AKA-print Aps

Forlaget Politica
c/o Institut for Statskundskab
Aarhus Universitet
8000 Århus C

Bogen udgives med støtte fra
Statens samfundsvidenskabelige Forskningsråd

PREFACE

This thesis should be considered as an attempt to integrate the methodological findings, arrived at in my work on the Danish election project at the Institute of Political Science since the 1970s, with special emphasis on the problem of ecological inference.

A number of colleagues and student research assistants have been very helpful during the (too) long research process. Especially I want to thank Jørgen Elklit for his encouragement and Jørgen Kirkegaard, Dan S. Hansen, Villads Villadsen and Orla Madsen for their keen and patient work with the overwhelming number of datafiles involved in the project.

I am grateful to Michael Væth from Mathematical Institute, University of Aarhus, for helpful comments on some of the mathematical derivations in chapter 3. I take, however, full responsibility for the application of the derivations to ecological inference.

It is a pleasure to thank my father Erik Risbjerg Thomsen, and colleague Noel Kent, for their attempts to correct my shaky English, and Ellen Bidstrup and Anette Riber Hansen for dealing so efficiently with a difficult manuscript.

It should be noticed that the numbering of equations, figures and tables starts over again in each chapter, and that the number of each chapter is indicated on the top of each page.

Aarhus, March 1987

Søren Risbjerg Thomsen

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Imagine

Imagine the following hypothetical situation:

At the same time as the proportional representation system was introduced at Danish parliamentary elections in 1920, a special agency was allowed to record the voting behaviour of every single voter at all ensuing elections.

To protect the secrecy of the vote it was not permitted to release information about the individual voter, but the agency was allowed to publish tables showing the individual mobility between the political parties from one election to the next as percentages of the total electorate. Furthermore, the agency could publish tables showing the party-distribution of the votes within the different social classes. - What material this would be for today's political science!

In fact, such a collection of tables actually exists, as displayed in appendices 2 and 3 to this book. However, I must at once admit that my "agency" presumably is not as reliable as the before-mentioned hypothetical agency would be, since the tables are output from a computer, only fed with official geographical election statistics, geographical census statistics, and with a computer program for what is called "ecological inference".

Beyond Sheer Mapping

To most students of electoral behaviour official election statistics are not very interesting. Especially since the post-war proliferation of political survey research, offering intimate descriptions of the ways and doings of the individual voters, the scientific study of geographic electoral data, called "electoral ecology", has been limited and marginal.

The neglect of electoral ecology has been justified by strong methodological arguments. In fact, the conclusion from the discussion has been, that it is still an open question if electoral ecology can contribute to little more than sheer geographical mapping of voting behaviour.

This is especially so when it comes to "ecological inference", i.e. estimation of individual behaviour from ecological covariation, such as ecological estimation of individual voter-mobility and class-voting (cf. chapter 3, section 2).

From time to time political researchers have ventured the dangerous journey beyond sheer geographical mapping, but such attempts have frequently been met by severe criticism.

It shall not be concealed that the present author is one of those researchers who - in spite of the criticism - think that the journey can be rewarding. In fact, although this book is about electoral ecology, not a single map shall be presented! Instead, the reader is invited to the presentation of a comprehensive methodological theory for aggregate election statistics which presumably can be used as an instrument for explaining voting behaviour as well at the aggregate as at the individual level.

The Inspiration from Rasch

As a sociology student at the University of Copenhagen in the late 1960s I had the luck that my professor in applied statistics, the late Georg Rasch, was an exceptional scholar in his field. He combined an academic background as a mathematician with an extensive experience as a statistical consultant in biomedical and psychological research. His enthusiastic and stimulating teachings which combined criticism towards established methods in applied statistics with ingenious proposals of alternative methods became a real fascination for some of the students. Especially his theory of measurement based on the almost philosophical concept of "specific objectivity" (cf. appendix 1) had high - even emotional - appeal in the small sect of "Rasch-fans" among the students. Given the fact that most social science students find statistical methods dusty and boring, it was a strange phenomenon, but I think that anyone who has had a similar experience with a gifted and enthusiastic teacher will know what I mean.

Rasch and his collaborators worked almost exclusively with individual-level data inside the field of psychometrics. But in my own work as a student and later on as a lecturer at the Institute of Political Science in Århus I tried to apply the Rasch theory of measurement to the analysis of aggregate political data.

The conclusion from my own work is that although the measurement models proposed by Rasch are too restrictive for the analysis of aggregate political data, they are very useful for the sociological interpretation of results obtained by more complex models. Hence although the methodological theory for electoral ecology presented in this book is not a straightforward application of the Rasch theory of measurement, the thinking of Georg Rasch has been a main inspiration.

The Rasch Models of Measurement

Given the vast amount of data in the social sciences based on qualitative responses, the problem for Rasch was to construct probability models for these qualitative data, which allowed quantitative measurement. Rasch shared this endeavour with other model-builders in contemporary social science, but he differed from these contemporaries by emphatically stressing the importance of objectivity in measurement. Regarding qualitative responses as produced by interaction between (social) objects and (social) stimuli he defined "specific objectivity" (or "separability") as comparison of objects independent of stimuli and comparison of stimuli independent of objects. By mathematical deduction from this very restrictive definition of objectivity Rasch succeeded in deriving a class of models, called "models of measurement" the validity of which is the necessary and sufficient prerequisite for objective comparison in the before mentioned restricted sense.The models are presented in appendix 1.

Rasch and collaborators demonstrated some very elegant applications of the models of measurement in the context of intelligence-, attainment- and attitude tests (Rasch, 1960, 1966; Andersen, 1966; Stene, 1968), but for some followers of Rasch the models of measurement constituted nearly an inescapable Procrustes bed that too many kinds of data had to be squeezed into.

The main problem with the models of measurement is simply that they often do not provide a very good description of actual data. To some "Rasch-fanatics" this would mean that the data are "no good", because they don't offer the opportunity for objective measurement. Another - more rational - reaction would be to throw away the Rasch-models and search for other models, which can provide a better description of the actual data.

My own solution to the problem is to regard the models of measurement as a base-line or null-hypothesis from which the description obtained with other models can be interpreted.

Logit Models

In the perspective of contemporary methodological theory for qualitative data, the class of measurement models, derived by G. Rasch, is only a special case of a more general class of models called "linear logit models" (chapter 2, section 2). These models have in recent years proved to be very good instruments for description of qualitative data within many branches of social science.

In this book I shall try to demonstrate that logit models also provide a very good description of electoral data. The problem is, however, that it is very difficult to give substantial interpretations of the results obtained by application of logit models. As implied above my solution to this problem is to use the measurement models of Rasch as a yardstick from which the results obtained by the logit models can be judged (chapter 2, section 2.2).

The Data

The methodology for electoral ecology proposed in this study is intended to be a general methodology, which can be used to describe electoral behaviour in different countries with different election systems. As material for demonstration of the applicability of the methodology I shall, however, mainly use official statistics from Danish parliamentary elections and censuses since 1920. These are the data I have worked with in "the Danish election project" at the Institute of Political Science applying the Danish statistical Commune Data Archive (Bentzon, 1975, 1985; Madsen, 1986). This database offers the opportunity to investigate the change in electoral behaviour in 246 constant geographical units throughout the period 1920-79.

To demonstrate the more general applicability of the methodology a limited amount of British electoral data shall also be analysed.

Special emphasis will be placed on analysis of across-time data because it is particularly in dynamic analysis that the strength of the methodology is obvious.

Analysis versus Inference

A very clear distinction shall be made between two kinds of methodologies, the methodology for ecological analysis and the methodology for ecological inference respectively.

The methodology for ecological analysis, described in chapter 2, has the objective of finding data-structures exclusively on the aggregate level. The problem here is to detect the pattern in the aggregate behaviour and to find causes for that behaviour on the aggregate level.

In contradistinction to that, the methodology for ecological inference, described in chapter 3, has not the objective of finding data-structures on the aggregate level but to construct "data" on the individual level from data on the aggregate ecological level. The methodology for ecological inference is similar to the methodology for forecasting, only the forecasting is not across-time but across-level.

Since the methodology for ecological analysis developed by the present author has extensively been described elsewhere (Thomsen, 1971, 1972, 1979) the main subject of this book is an attempt to solve the problem of ecological inference. However, this attempt draws heavily on the experience from my work with ecological analysis.

Macro versus Micro Interpretation

Parallel to the distinction between the methodology for ecological analysis and the methodology for ecological inference is the distinction between two kinds of substantial interpretations, i.e. macro versus micro interpretation. On the one hand the macro-interpretation is the "genuine" sociological interpretation where all concepts used are defined for organized aggregates, e.g. local communities or social classes. On the other hand the micro-interpretation is the social-psychological interpretation which includes concepts about individual reactions.

Ecological Analysis

In the kind of electoral ecological analysis presented in this book (chapter 2) the units of analysis are populations of voters within geographical areas such as constituencies or election districts (i.e. subunits within constituencies) and the variables characterizing the units are proportions of voters in different categories. In the Danish election project the party-distribution (including the "party" of non-voting) as well as the social (e.g. class, occupational, age) distribution of the voters are recorded for each election district.

Now the problem in carrying out the ecological analysis is then to describe and explain the statistical variation of the party-distributions across districts and across elections. This is done by constructing mathematical models, supposed to reflect the empirical structure of the ecological data, and then confronting the models with the data.

If a model is considered a valid description of the actual data, it can be used for a substantial political sociological interpretation of the recorded electoral behaviour. The sociological interpretation can be eased by including data about the social structure of the ecological units in the analysis.

An important purpose of the present work is to demonstrate that it is not sufficient to construct models that provide good statistical descriptions of actual data, the models should also be well suited for substantial interpretations.

Ecological Inference

While only macro-interpretations are offered in connection with ecological analysis, ecological inference can be considered a daring statistical technique for making micro-interpretations from ecological data.

Since the 1950s scientists have tried to solve the problem of ecological inference but, as discussed in chapter 3, they have not been very succesful within electoral research. The most promising attempt to my knowledge is the ecological estimation of the British two-party vote within two classes (cf. chapter 3, section 2).

In my opinion it seems that the methodology, mainly based on regression techniques, developed to this day, fails when the ecological estimation concerns more than two parties or more than two classes. In this perspective the work presented in this book is an attempt to develop a new methodology based on latent structure theory for the ecological estimation of individual voting behaviour concerning many parties and many classes (in the case of Denmark: more than 10 parties, and up to 7 classes). The validity of the results is discussed in chapter 4.

Ecological Analysis of Results from Ecological Inference

The final subject in this book is an integration of the two methodologies presented. It is the ecological analysis of results obtained by ecological inference. Since social classes just as well as election districts are aggregates characterized by party-distributions, the methodology for ecological analysis can be applied to the "data" constructed for each class by the methodology for ecological inference.

In this context as index for the amount of "class-polarisation" is presented in appendix 1, section 3, and applied in chapter 4, section 3, in connection with the analysis of electoral behaviour in Denmark 1920-79.

CHAPTER 2

ECOLOGICAL ANALYSIS

This chapter presents a methodology for the analysis of change in electoral ecological behaviour. The term "ecological" means, as already mentioned, that the units of analysis are populations within geographical areas such as constituencies or election districts. The kind of electoral behaviour under study is typically the support for different political parties in the different ecological units, where the support for each party in each unit is considered as the party's relative share of all voters in the unit.[1]

The methodology primarily aims at kinds of elections, such as general parliamentary elections or presidental elections, giving the voters the same choice-alternatives, i.e. it is possible to choose between the same parties or between the same presidental candidates in every unit. It should at least be possible to subdivide the ecological units into fairly large groups of units (regions) with the same choice-alternatives for all voters.

The methodology will mainly be illustrated by across-time analysis of Danish parliamentary elections, but although the methodology was developed within the Danish context, I believe it has broad applicability for other electoral systems as well. To demonstrate the broader applicability of the methodology an example from British general elections is presented.

The methodological approach is based on an application of exponential mathematical probability models. Within this approach three models for change in ecological voting behaviour are presented. In order of presentation the three models are:

1. A model for uniform ecological swing which applies to small as well as big parties
2. An ecological linear logit model for explaining deviations from uniform swing
3. An ecological multiple choice model for multi-party systems.

The first model is in fact a special case of the second model, and the second model a special case of the third model. As follows in chapter 3, these three ecological models are incorporated into a comprehensive methodological theory which includes voting behaviour at the individual level as well, offering the opportunity of ecological inference from the aggregate level to the individual level.

The presentation in this chapter is inductive, i.e. the models are formulated through several steps of generalization as data-structures of increasing complexity are being introduced.

Strictly speaking the two first models presented apply to genuine binary choice only, so the voter has two and only two possible alternatives at each election. This constraint is totally unrealistic, because even in two-party systems the third alternative of "non-voting" is a possible choice. The methodology can, however, with some approximation be applied to analyse "crude binary choice" which means that one party or group of parties is considered as the first alternative while the "other" parties (incl. the "party" of non-voting) are considered as the second alternative.

1 A Model for Uniform Swing

An example of the kind of data needed to test the model for uniform swing appears in table 1. It shows the proportion of all voters voting for the Social Democratic Party in two districts (constituencies) at the Danish general elections in 1971 and 1973.

Table 1.1 Support (proportions of voters) for Social Democrats in 1971 and 1973 in two districts[1]

	Support		Change in support		
District no.	1971	1973	Difference of support	Ratio of support	Ratio of odds
i	PX_i	PY_i	PY_i-PX_i	PY_i/PX_i	$\frac{PY_i/(1-PY_i)}{PX_i/(1-PX_i)}$
1	.445	.331	-.114	.74	.62
2	.209	.137	-.072	.66	.60

1) District no. 1 is Randers Constituency and district no. 2 is Varde Constituency.

The support for the Social Democrats went down in district no. 1 from 44.5 per cent to 33.1 per cent of the voters and in district no. 2 from 20.9 to 13.7 per cent. The 1973 election was a defeat for the Social Democrats unprecedented since 1945 and the party lost support in all 103 constituencies. But was the decrease in support of the same size in all districts? Clearly the answer to this question depends on how we measure the change in support from one election to the next.

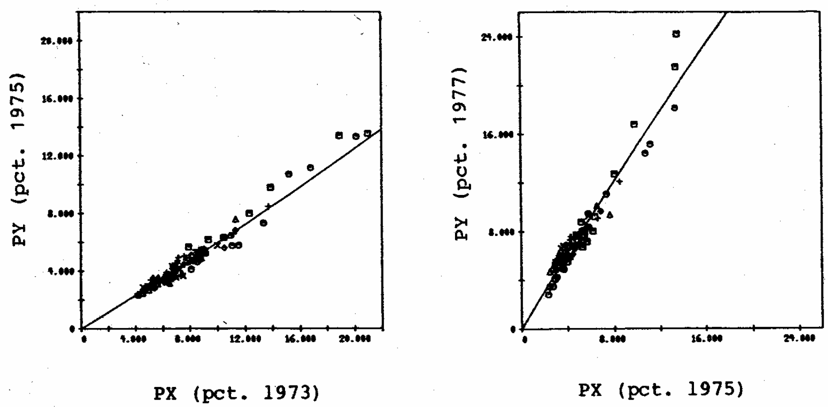
Probably the most common way to measure change in support is to subtract the support at the first election from the support at the second election as indicated in column 3 of table 1. Using this measure for change clearly the loss of support was greater in district 1 (-11.4 pct.) than in district 2 (-7.2 pct.). The conclusion is that the loss of support was not a "uniform swing" for the Social Democrats and we can start looking for explanations why the party in spite of the defeat managed better in district 2 than in district 1. But before doing that it may be worthwhile to speculate if the implicit model behind the calculation of change is appropriate. The implicit model for uniform "additive" change is simply (using the notation from table 1.1)

$$PY_i - PX_i = c \quad ; \quad i=1,2,\ldots,n \qquad (1.1)$$

that is, the difference of support is equal to a constant c independent of the individual district (n is the number of districts).

One problem with the "constant difference of support model" is that a party cannot possibly lose more votes than it has already got, hence the negative constant c in formula (1.1) cannot be greater than the weakest support in any district at the first election. Because of this "floor effect" uniform swing is rather unlikely in political systems with many small parties participating in all districts like in the proportional representation systems in the Scandinavian nations. A closer look at election results in Scandinavia for small parties in fact reveals a typical "multiplicative" pattern, that is, the party support at the second election tends to be proportional to the support at the first election in districts all over the country (cf. Borre and Stehouwer, 1970; Gustafsson, 1974). Two examples of multiplicative change appears in figure 1.1. The graph (a) shows the relation between the support in 1973 and the support in 1975 for the Conservative Party in all Danish constituencies. The party lost votes from 1973 to 1975 but gained votes from 1975 to 1977 which appears in graph (b). In both graphs the points are fitted with a straight line through the origin which indicates that the second election support is proportional to the first election support.

Returning to table 1.1, column 4 shows the "multiplicative increase" that is, the ratio of support in 1973 for the Social Democratic Party to the support in 1971 in the two districts. With this measure the conclusion is quite the opposite from the previous one. Now the decrease is greater (the increase ratio is lower: .66) in district 2 than in district 1 (increase ratio: .74), and it must now be explained what



(a) (b)

Figure 1.1 Support for the Danish Conservative Party (a): 1973-1975, (b): 1975-1977. Constituencies

went especially wrong for the Social Democrats in district 2 compared to district 1! But wait - let's turn to the implicit model behind the calculations. The implicit model for "multiplicative" uniform change is

$$\frac{PY_i}{PX_i} = c \quad ; \; i=1,2,\ldots,n \qquad (1.2)$$

that is, the ratio of support is the same in all districts. This "constant ratio of support model" is very useful for small parties, but for larger parties it leads to absurd results. For instance if the support for a party in one district increases from 10 to 20 pct., it should, according to the "constant ratio of support model" in another district with an in initial support of 60 pct. increase to 120 pct.! This is of course not logically possible, and a "ceiling-effect" is observed the closer the support comes to 100 pct. In fact it seems that the "constant difference of support model", formula (1.1), is a better model for parties having a national support over 30 pct. and it was proposed by Stokes (1969) for description of the swing at the British general elections[2].

Of course it is most unsatisfactory to use different models depending on the size of the party involved. Fortunately a model can be formulated which has the attractive properties that it exhibits nearly additive change for big parties and nearly multiplicative change for small parties. It is the "constant odds ratio model" which assumes that the multiplicative increase in the odds for a certain party (odds is the proportion voting for the party divided by the proportion not voting for the party) is constant across districts[3]:

$$\frac{PY_i/(1-PY_i)}{PX_i/(1-PX_i)} = c \quad ; \; i=1,2,\ldots,n \qquad (1.3)$$

Column 5 in table 1.1 shows that the odds ratio between 1973 and 1971 is virtually the same in the two districts. With this model the swing for the Social Democrats is nearly uniform and no explanation specific to the individual district is needed.

For the constant odds ratio model the functional relation between PX and PY derived from formula (1.3) is (cf. Thomsen, 1979, pp. 90-91)

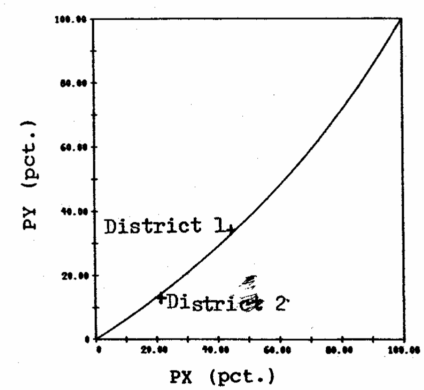


Figure 1.2 The relation between PX and PY according to the constant odds ratio model. c = .61

$$PY = \frac{cPX}{1+(c-1)PX} \qquad (1.4)$$

and figure 1.2 shows this relation for c = .61 (which is the average odds ratio for the two districts in table 1.1). The SD-support in the two districts 1971 to 1973 is plotted in the graph.

Furthermore in figure 1.3 (after rescaling) the same curve together with the SD-support in all 103 Danish constituencies is plotted with special signatures for different geographical regions. The plot indicates that for the individual regions a closer fit to the curve would occur with a special curve for each region. For instance it appears that the districts in the Copenhagen region (KBNHAVN) could be fitted with a curve for uniform change which has a higher ratio of odds than the other regions.

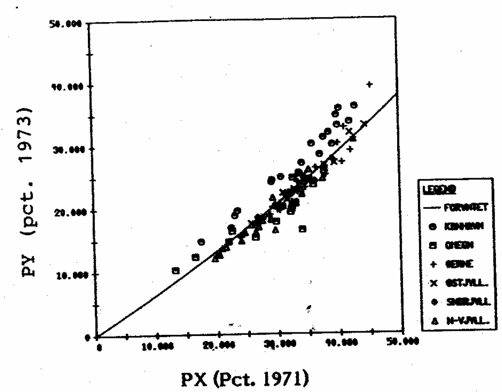


Figure 1.3 Support for the Danish Social Democratic Party in all Constituencies 1971-73. Curve for Constant Odds Ratio, c = .61

Figure 1.4 shows different curves for constant odds ratio with different values of c as indicated in the graph. The individual curves show characteristic floor and ceiling effects for either very low or very high proportions, and can for low proportions be approximated by straight lines through the origin, i.e. multiplicative change. For proportions close to 50 pct. and moderate values of c (.5 < c < 2.) the curves are nearly parallel to the diagonal, i.e. indication of nearly additive change.

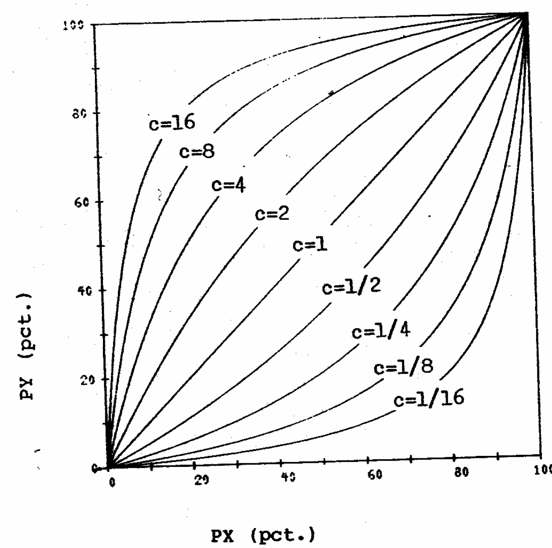


Figure 1.4 Different Curves for Constant Odds Ratio with Different Values of c as Indicated

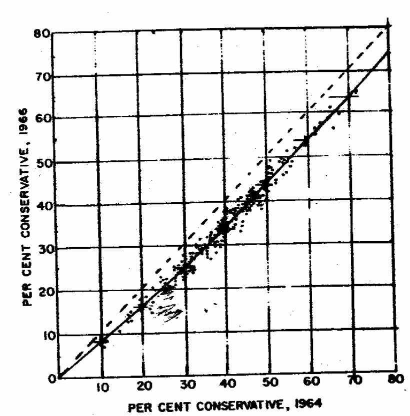


Figure 1.5 British General Elections. Conservative Share of Vote in 1964 and 1966 in Seats Contested by only Conservatives and Labour. Curve for Constant Odds Ratio, c=.77

Figure 1.5 shows, that the swing at the British general elections from 1964 to 1966 can be fitted with the constant odds ratio curve with c=.77. It is the same data that Stokes' (1969) described with a constant proportion difference model. Even though Stokes additive model is more simple than the odds ratio model and probably minimizes the unexplained variance quite as well, the advantage of the odds ratio model is, that it has greater validity in comparative research involving different political systems with both small and big parties.

In the next section the constant odds ratio model is developed as a special case of a general linear logit model for ecological voting behaviour. This special case can be interpreted as a situation where the "social orientation" of the party (e.g. the class-cleavage position) is unchanged from the first to the second election, while the "general popularity" of the party may change, hence the uniform swing. I think that the broad applicability of the ratio of odds model rests on the fact that in many political systems the social orientation of the parties is very stable over time while the general popularity is easily influenced by short-term factors such as the achievement of the governing party, leader popularity, party-position on variating national issues etc.[4]

Normally the range of variation of the support for a certain party (across districts) is limited so that the simple linear proportion model

$$PY_i = \alpha+\beta PX_i \quad ; \; i=1,2,\ldots,n \tag{1.5}$$

is a reasonal approximation to the constant odds ratio model. Although the simple model (1.5) describes the covariation between PX and PY just as well as the constant ratio of odds model (1.4), the problem concerning the simple model is, that the interpretation of the parameters α and β is very complicated. In fact, if the constant odds ratio model is valid, both α and β are complex functions of both the odds ratio c and the average support for the party across districts (cf. Thomsen, 1979, pp. 91-95).

2. The Linear Logit Model

In the last two decades social science methodologists working with qualitative data have become increasingly aware that it is unsatisfactory to use linear models for direct description of proportions (or probabilities) of individual qualitative behaviour (such as responses to psychological tests, choice of transportation, etc.)[5] One important difficulty with the direct linear approach is, that proportions vary

beween 0 and 1 while the implicit interval-scale assumption in linear models is that the range of variation is infinite. A common solution to this problem is to assume that the proportion PX of a certain qualitative alternative X is a function of an interval-scale property LX with S-curve shape like in figure 6.

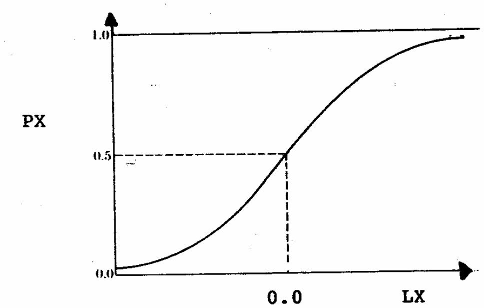


Figure 2.1 -The S-Curve Relation between LX and PX

An example of such a S-curve function is the logistic function

$$PX = \Psi(LX) = \frac{\exp(LX)}{1+\exp(LX)} \tag{2.1}$$

where exp(LX) is the exponential transformation of LX. LX can be computed from PX by

$$LX = \Psi^{-1}(PX) = \ln\left[\frac{PX}{1-PX}\right] \tag{2.2}$$

i.e. the natural logarithm to odds for the party X. This transformation of PX is known as the "logit" transformation (analogous to the "probit" transformation when the cumulative normal distribution function is used) (Theil, 1966, p. 632). Besides the infinite variation of LX it has the attractive property that it gives perfect symmetric results for non-occurrence of X (non X: $\bar{X}$), i.e.

$$LX = -L\bar{X} \qquad (2.3)$$

A model for explaining the variation of the "logit-support" LX is "the linear logit model":[6]

$$LX = a_o + a\mu + e \qquad (2.4)$$

where μ is a column vector of variables

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_k \end{pmatrix}$$

variating across disctricts and indicating each districts position on each of the k social cleavage-dimensions. a is a row-vector of parameters

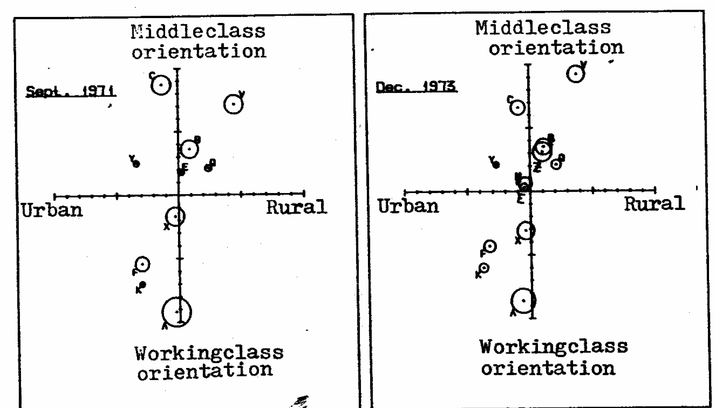
$$a = (a_1, a_2, \ldots, a_k)$$

specific to the party and indicating the party's social orientation on each dimension, while a_o is the general popularity of the party independent of the social orientation. e is a random variable with zero mean describing the variation of LX not explained by μ. For the district no. i the model can be written

$$LX_i = a_o + a_1\mu_{1i} + a_2\mu_{2i} + \ldots + a_k\mu_{ki} + e_i \qquad (2.5)$$

The parameters of the model can be estimated by factor analysis or regression methods. When factor analytical methods are applied, the logit-support for the different parties at different elections are taken as data, and the "factor-weights" $a_1, a_2, \ldots, a_k$ are estimated by raising assumptions about the number k of dimensions and the population distribution of the "factors" $\mu_1, \mu_2, \ldots, \mu_k$ (Harman, 1967; Jackson et al., 1979). Normally, it is assumed that each factor is standard normal distributed and that the factors are orthogonal (i.e. uncorrelated). The standard problem in connection with factor analysis is to interpret the estimated factors, but this problem can be eased by including a number of structural variables in the analysis.

In the Danish election project the two most important social cleavage dimensions were identified by factor analysis as the rural-urban dimension and the middleclass-workingclass dimension throughout the period 1920-1979 (Thomsen, 1979). It turned out that the two dimensions could be measured by logit transformations of the proportion of voters in agriculture and of the proportion of middleclass voters, and the social orientation of each party was estimated as regression coefficients to these two variables. An interesting result was, that the regression coefficients were much more stable when logits were used instead of



A: Social Democrats
B: Radical Liberals
C: Conservatives
E: Justice Party
F: Socialist Peopl.P.
K: Communists
M: Centre Democrats
Q: Christian Peopl.P.
V: Agrarian Liberals
X: Non-voters
Y: Left Socialists
Z: Progress Party

Figure 2.2 Social Orientation of Political Parties in Denmark Estimated by Ecological Regression Method 1971 and 1973 proportions as well for the dependent logit-support variable (especially for small parties) as for the independent variables. This result indicates that the linear logit model has greater validity than the ordinary linear proportion model often used in political ecological research (e.g. Janson, 1969). Another interesting result was that the most important factor in the factor analysis (i.e. the factor which explained most of the variance in the data) was the rural-urban dimension, while the most important regressor in the regression analysis (i.e. the regressor with highest coefficients) was the middleclass-workingclass variable. This result is explained by the important fact that the rural-urban variable in the ecological data has much greater variance than the middleclass-workingclass variable. In other words it is as if the "vertical" class-dimension in ecological data is "concealed" by the "horizontal" rural-urban dimension.

Results from the logit regression analysis for 1971 and 1973 are presented in figure 2.2. The horizontal axis indicates the party's position (a_1) on the rural-urban dimension, and the vertical axis is the position (a_2) on the middleclass-workingclass dimension. Each party is represented by the center of a circle and the area of the circle is proportional to the national support for the party. The figure shows for instance that the social orientation of the Social Democratic party is virtually unchanged from 1971 to 1973, while this is not so for the Agrarian Liberals.

2.1 Dymamic Properties of the Linear Logit Model

This subsection describes some short-term dynamic properties of the linear logit model by considering two successive elections close in time. Because the social structure of the districts only change slowly we assume that μ (the district's social position) is approximately constant over time while the general popularity and the social orientation of the party may change, hence the following model for logit-support LX at election 1 and logit-support LY at election 2 for a certain party in district no. i:

$$LX_i = a_o + a\mu_i + e_{1i}$$
$$LY_i = b_o + b\mu_i + e_{2i} \qquad (2.6)$$

a_o and b_o are the general popularity, a and b are the social orientation, e_{1i} and e_{2i} are the residual random components for a certain party at election 1 and election 2 respectively.

For the logit-support the change over time is supposed to be additive, consequently the amount of change in a district is measured by subtracting LX_i from LY_i:

$$LY_i - LX_i = (b_o - a_o) + (b-a)\mu_i + (e_{2i} - e_{1i}) \qquad (2.7)$$

Because a political party for historical reasons often has "political strongholds" independent of the current social structure the residuals e_{1i} and e_{2i} from the linear structure normally show autocorrelation by having positive covariance (across districts), i.e.

$$\mathrm{Cov}(e_1, e_2) > 0 \qquad (2.8)$$

And from this fact follows that the variance (across districts) of the difference between the two residuals is

$$\mathrm{Var}(e_2 - e_1) = \mathrm{Var}(e_1) + \mathrm{Var}(e_2) - 2\mathrm{Cov}(e_1, e_2)$$
$$< \mathrm{Var}(e_1) + \mathrm{Var}(e_2) \qquad (2.9)$$

that is, the variance of the difference between the residuals is less than the sum of the variance on each residual. Furthermore the covariance between the residuals are often so strong that the variance of the difference is negligible. So for simplicity of the following discussions we shall assume that a close approximation to formula (2.7) is

$$LY_i - LX_i = (b_o - a_o) + (b-a)\mu_i \qquad (2.10)$$

where the random component is omitted.

As mentioned in the previous sections the social orientation of the parties is often very stable while the general popularity may change. The special case of unchanged social orientation is defined

$$a = b \qquad (2.11)$$

and consequently in this special case of unchanged social orientation formula (2.10) becomes

$$LY_i - LX_i = b_o - a_o \qquad (2.12)$$

that is, the difference in the logit-support is independent of the districts.

Substituting PX and PY for LX and LY by equation (2.2) the model (2.13) for "constant social orientation" becomes

$$\frac{PY_i/(1-PY_i)}{PX_i/(1-PX_i)} = \exp(b_o - a_o) \qquad (2.13)$$

This model is exactly the same as the constant odds ratio model for uniform swing, equation (1.3), presented in section 1, and the constant c in equation (1.3) can now be interpreted as the exponential transformation of the increase in general popularity defined by the linear logit model.

2.2 Interpretation in Relation to the Rasch Theory of Separability

The model (2.12) for constant social orientation can be considered as the aggregate version of the Rasch model of measurement for binary choice (cf. appendix 1). As discussed in appendix 1 the validity of this model allows for "separability" which, when applied to electoral behaviour, means that the covariation between parties and elections is statistically independent of the districts, and vice versa, the

covariation between parties and districts is statistically independent of the elections. Furthermore, it has been proved by Rasch (1968) that it is also the only model for binary choice, which allows for separability.

This important result from mathematical statistics places the model (2.12) in a unique position, which has been largely overlooked by social scientists. For this writer the measurement theory of Rasch makes it very difficult to conceive of a probability model which better than the model (2.12) represents the situation where change in the support for a party does not interact with the attributes of the districts, and hence can be interpreted as a model for unchanged social orientation of the party.

2.3. Deviations from Uniform Swing

Returning to the general model, equation (2.10), which allows for change in social orientation, the term $(b-a)\mu_i$ can now be interpreted as a model for explaining the deviations from uniform swing as defined by formula (1.3). These deviations can be investigated by factor analysis or regression methods.

In the special case where μ is one-dimensional, i.e. only one dimension describes the variation in the logit-support for a certain party, the functional relation between LY and LX is

$$LY_i = \left[b_o - \frac{b}{a}a_o\right] + \frac{b}{a}LX_i \qquad (2.14)$$

i.e. a simple linear relation holds between LY and LX. As for the constant ratio of odds model normally the range of variation for PX and PY is limited so that the simple linear proportion model

$$PY_i = \alpha + \beta PX_i \qquad (2.15)$$

is a reasonable approximation to equation (2.14). Even though the correlation between PX and PY often is about the same as the correlation between LX and LY again the problem is, that the interpretation of the parameters α and β is very complicated in relation to the linear logit model. This conception has important implications for a number of previous dynamic electoral ecological studies which have employed the simple linear proportion model (2.15) (e.g. Borre and Stehouwer, 1970; Gustafsson, 1974; Hoschka and Schunck, 1978). If my conception is valid most of the interpretations arrived at in these studies in connection with the linear proportion model become ambiguous.

Figure 2.3 shows an example where the higher validity of the model (2.14) compared to the model (2.15) is obvious.

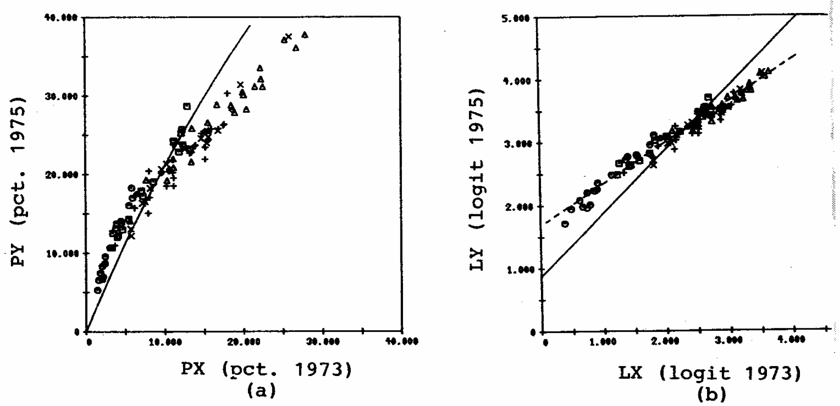


Figure 2.3 Support for Agrarian Liberal Party 1973-1975. (a): Proportions; (b): Logits

The two graphs show the relation between PX and PY as well as LX and LY for the Agrarian Liberals 1973 to 1975. It is the period where the relatively small party headed by a strong party leader seized power by forming a single party minority government which forced tax-cuts and proposed an austerity programme before appealing to the country in 1975. The graph (a) shows strong deviations from the constant ratio of odds curve in a non-linear way but the relation between LX and LY in graph (b) is roughly linear with a slope less than 1 (the solid straight line in graph (b) with slope equal 1 is the constant ratio of odds curve after logit-transformation). According to formula (2.14) the linear relation in graph (b) is expected between LX and LY, if the one-dimensional linear logit model is valid. Graph (b) shows that the Agrarian liberals changed social orientation to a more moderate positon on a dimension which is highly correlated to their support in 1973.

3 The Linear Multiple Logit Model

Strictly speaking, the models discussed until now applies only to genuine binary choice, i.e. the voter has two and only two possible alternatives at each election. As mentioned in the introduction to this chapter this situation is not realistic, because even in two-party systems the third option of non-voting is a possible choice. This section formulates a model for multiple choice called "the linear multiple logit model" (the LML model), which can be regarded as a generalization of the linear logit model (the LL model). It was originally proposed by G. Rasch (1961) and its application to electoral research was described by the present author (Thomsen, 1971, 1972). The model was "reinvented" in econometrics by Theil (1969); see also Aldrich and Nelson (1984, pp. 37-40).

The main difficulty in using models for binary choice in multiple choice situations is, that these models do not account for the complex interaction between all possible options available to the individual. So in multi-party choice situations the simple S-curve relation, equation (2.1), between the proportion PX voting for a party and the interval-scale property LX is not very likely. However, a simple generalization of equation (2.1) is

$$PX(h) = \frac{\exp[MX(h)]}{\exp[MX(1)]+\exp[MX(2)]+...+\exp[MX(m)]} \quad (3.1)$$

where PX(h) is the proportion choosing party no. h out of all m parties and MX(h) is the socalled multiple logit interval-scale property for choosing party no. h. Equation (2.1) can be regarded as a special case of equation (3.1) for m = 2 by the definitions:

$$PX = PX(1) \quad ; \quad 1-PX = PX(2) \quad ; \quad LX = MX(1)-MX(2) \quad (3.2)$$

MX(h) can be regarded as the "strength" of party no. h in relation to all other parties and equation (3.1) accounts for the "competition" among the parties by setting the strength of one party in relation to the strength of all the parties competing in a district.

Analogous with equation (2.4) the LML model is

$$MX(h) = a_o(h) + a(h)\mu + e(h) \qquad (3.3)$$

where μ has the same interpretation as for the LL model, equation (2.4), i.e. the districts position on a number of social cleavage dimensions, and $a_o(h)$, $a(h)$, $e(h)$ are the general popularity, the social orientation, the random component for party no. h independent of the districts. Similar to the LL model the parameters of the LML model can be estimated by factor analysis or regression methods.

The LL model, eq. (2.4), can be regarded as a special case of the LML model for m = 2 by the definitions (3.2) and the definitions:

$$a_o = a_o(1) - a_o(2) \;\; ; \;\; a = a(1) - a(2) \;\; ; \;\; e = e(1) - e(2) \qquad (3.4)$$

In multiple choice situations this result means that it is in accordance with the LML model to use the LL model by paired comparison of the parties as done by Miller (1981).

For "crude binary choice" (i.e. one party or group of parties are considered as the first alternative while the other parties are considered the second alternative) in multiple choice situations the mathematical analysis of the relation between the LML-model and the LL-model is complicated as presented in chapter 3, subsection 4.2. A main result is, that the LL-model is only a crude approximation to the LML-model in this situation. Furthermore, the constant ratio of odds model for uniform swing is only a crude approximation to the special case of the LML-model with constant social orientation across time (i.e. the multiple choice model for uniform swing, discussed in appendix 1).

Bearing this result in mind it is a surprising and an important finding from the Danish election project that the LML model does not give a better description of the electoral data than the LL model or a different conclusion about the development of the social orientation of the parties, even though at some elections more than 10 parties participate in all electoral districts. This indicates, that the LL model is a reasonable approximation to the LML model for this kind of data, at least in the Danish case with many small parties.

Although the LML Model is not especially useful for ecological analysis, this is not the case when the level of analysis is the individual voters. This is one of the themes in the next chapter.

Notes to Chapter 2

1. In electoral ecological research it is very common to consider the support for each party as the share of all valid votes cast (excluding the usually very small "party" of invalid votes and the usually not insignificant "party" of non-voting). This choice is justified if the interest of the researcher primarily concerns the political consequences of the elections. However, if the aim of the analysis is the sociological explanation of the voters electoral behaviour in general, it is a better choice to consider the support for each "party" as its share of all voters.

2. The problem of possible third parties will be treated later. To avoid the problem of third parties having candidates in some of the constituencies at the British general elections the swing is measured as the average gain for Conservatives and loss for Labour. (Butler and Stokes, 1971, p. 335).

3. The odds ratio model has been used extensively in many fields (cf. Fleiss, 1973; Plackett, 1974; Breslow and Day, 1980).

4. The constant odds ratio model has been used by the present author for election night forecasting within the theoretical framework of bayesian Statistics (Thomsen, 1981b).

5. For a recent introduction see Aldrich and Nelson (1984). The two most important schools of methodology within social science using non-linear probability models are the psychometricians and the econometricians.

6. It is called "the linear logit model" by the econometricians (Theil, 1966, p. 133). The psychometricians call it "the logistic test model" (Lord and Novick, 1968, pp. 397-479). It has been used in many other research fields as well (cf. Cox, 1970, chapter 2; Plackett, 1974; chapter 2-3; Breslow and Day, 1980, chapter 6).

CHAPTER 3

ECOLOGICAL INFERENCE

1 The Problem of Ecological Inference

The problem of ecological inference is simple enough to formulate: we want to use official aggregate statistics to say something about the behaviour of individuals. With regard to electoral research we want to answer questions about voter mobility like, "How many voters changed from party X to party Y from election 1 to election 2?" and about class voting, "How many voters in class A voted for party X?" without conducting expensive surveys. Especially in historical research the need for ecological inference is great, simply because we often do not have information on past individual behaviour.

An example of ecological covariation from Danish parliamentary elections is shown in figure 1.1.

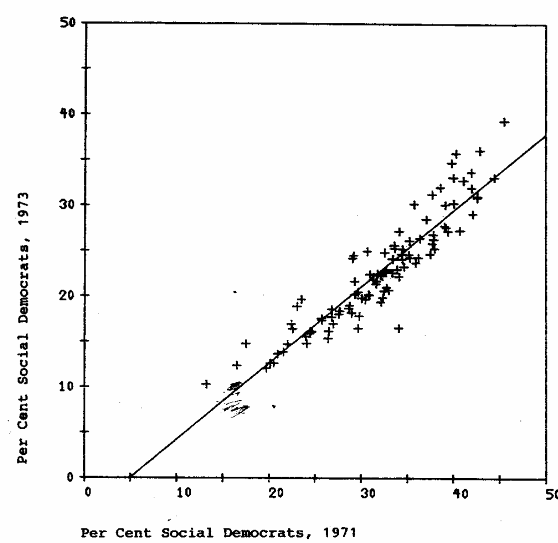


Fig. 1.1 Example of Ecological Covariation. Aggregate Voter Mobility for Constituencies at Danish Parliamentary Elections 1971-

Each point in the graph represents a constituency, and the projection on the horizontal X-axis is the percentage PX of all voters in the constituency voting for the Social Democrats at the election in 1971 while the projection on the vertical Y-axis is the percentage PY voting for the Social Democrats at the election in 1973. Probably because of the unprecedented success of two new parties, the Progress Party and the Centre Democrats, the 1973-election was a near catastrophe for the Social Democrats who suffered a severe setback in every constituency.

Anyhow figure 1.1 shows a strong positive covariation, i.e. high percentages in 1971 were followed by relatively high percentages in 1973, and vice versa low percentages were followed by relatively low percentages. The degree of covariation can be described either by the slope of the regression line plotted in figure 1.1 or by the fit to the regression line indicated by the correlation coefficient, which computes the square root of the proportion of variation in PY explained by the variation in PX.

Now the problem of ecological inference is to estimate the individual covariation shown in table 1.1 from the ecological covariation shown in figure 1.1.

Table 1.1 Example of Individual Covariation Obtained From Survey Results[1]. Voter Mobility for Individual Voters at Danish Parliamentary Elections 1971-1973. Per Cent of All Voters.

Election 1971	Election 1973: Social Democrats	Election 1973: Other parties*	Election 1973: Total
Social Democrats	20.3	12.0	32.3
Other Parties	2.3	65.4	67.7
Total	22.6	77.4	100.0

Note 1): Survey results (Worre, 1976 p. 21). The percentages have been adjusted so that the marginals of the table are the true marginals from the election results.

* Other parties include the "party" of non-voting.

Since the voting procedure is done in secret we don't know the content of table 1.1 unless the individual voters are interviewed. We only know the marginal distributions in table 1.1 from official election results. Anyhow it is clear that table 1.1 must have some relation to figure 1.1; it is the same behaviour, but described on different levels.

At least one should think so! The practice of infering individual-level behaviour from aggregate level behaviour is called ecological inference (e.g. Langbein and Lichtman, 1978) or disaggregation (e.g. Hannan, 1971). In 1950 this practice was very much discredited by a short article by Robinson, appearing in the American Sociological Review, where it was given the nickname "the ecological fallacy"; i.e. the false belief that ecological inference is possible. Robinson demonstrated his point of view with a number of examples showing that the Pearson product-moment correlation coefficient between two variables on the individual level highly differed from the corresponding correlation between the average values of the two variables computed for a number of geographical areas.

Robinson's article had (and still has) a strong impact on methodological thinking on problems of ecological inference in particular, and problems of cross-level inference in general (e.g. Galtung, 1967, p. 47; Alker, 1969).

The rationale behind the criticism on ecological inference is simply that there is no necessary logical connection between individual variation and ecological variation (i.e. the variations across aggregates). More precisely there is no necessary logical connection between the within-aggregate individual variation and the between-aggregate variation, while the total individual variation can be considered as a mixture of within-aggregate and between-aggregate variation (cf. Cronbach, 1976; Firebaugh, 1978).

In my opinion, to take the logical gap between individual and ecological variation as an insuperable obstacle for ecological inference is just as futile as to take the logical gap between the present and the future as an insuperable obstacle for making predictions. What is needed instead is theoretical propositions and empirical studies concerning the relation between the individual and the ecological level. In fact, a rich research tradition to fulfil this need has developed in American sociology as a reaction against the pessimism inherent in the concept of "the ecological fallacy". Following a seminal article by Leo Goodman (1959) a lot of work has been done in the 1970's mainly concerning bivariate relationships (Hannan, 1977; Hammond, 1973; Iversen, 1973; Hannuchek et al., 1974; Hannan and Burstein, 1974; Smith, 1977; Burstein, 1978; Firebaugh, 1978). Excellent review articles are: Haney, 1980; Glick, 1980). In the 1980's the focus of interest has turned to multivariate relationships (Lincoln and Zeitz, 1980; Stipak and Hensler, 1982; Althauser et. al., 1982). A main conclusion from the quoted work is that regression techniques are much more appropriate than correlation techniques when establishing the relation between the individual and the ecological level.

2 Criticism of the Regression Approach

Unfortunately to this day electoral ecological analysts have not been very successful in establishing the relation between the individual and the ecological level. The most promising attempt to my knowledge are the ecological estimations of the British two-party vote within two classes, manual and non-manual made by Crewe and Payne (1976). The validity of ecological estimates of voter mobility seems to be considerably lower (Stokes, 1969; Hoschka and Schunck, 1975), especially for multi-party voter mobility. This is also the case for attempted ecological estimation of multi-party class voting in more than two classes, which can even lead to absurd results (Lewin et al., 1972).

Common to nearly all contemporary work with ecological inference in electoral research as well as in other branches of social science is what I shall call "the regression approach". By this approach individual behaviour is inferred from the results obtained by regression analysis carried out on the aggregate level.

As a starting point the methodology for ecological inference by the regression approach shall be illustrated by a procedure treated by a number of authors applying the simple regression model within electoral research.

2.1 Ecological Inference by the Simple Regression Model

The methodology for ecological inference by the simple regression model can be demonstrated by the attempted ecological estimation of two-party voter mobility between two elections (cf. Stokes, 1969).

At election No. 1 the two parties are the party X and the party $\bar{X}$ (non X), while at election No. 2 the two parties are the party Y and the party $\bar{Y}$ (non Y). The choice between X and $\bar{X}$ can also be considered as a "crude binary choice" between the party X and all other parties $\bar{X}$. The same methodology applies when X and $\bar{X}$ are two social classes and Y and $\bar{Y}$ are two political parties (cf. Hammond, 1973, p. 767-768; Crewe and Payne, 1976, pp. 47-48; Miller, 1977, p. 42). On the individual level the model behind the estimation (which often is not explicitly stated by the different authors) is:

$$PY_{ij} = P(Y|X)PX_{ij}+P(Y|\bar{X})P\bar{X}_{ij}+u_{ij} \; ; \quad \begin{matrix} i=1,2,\ldots,n \\ j=1,2,\ldots,N_i \end{matrix} \qquad (2.1)$$

where n is the number of election districts, and N_i is the number of voters in district No. i. PX_{ij} is the probability that the individual voter No. j in district No. i chooses party X at election 1 and PY_{ij} is the probability that the same voter chooses party Y at election 2. P(Y X) is the conditional probability for choosing Y at election 2, given that X has been chosen at election 1. u_{ij} is stochastic disturbance uncorrelated with PX_{ij}.

The crucial feature to eq. (2.1) is that the conditional probabilities P(Y X) and P(Y $\bar{X}$) are constant across all individuals in all districts. Since

$$P\bar{X}_{ij} = 1 - PX_{ij}$$

equation (2.1) becomes

$$\begin{aligned} PY_{ij} &= P(Y|X)PX_{ij}+P(Y|\bar{X})(1-PX_{ij})+u_{ij} \\ &= P(Y|\bar{X})+[P(Y|X)-P(Y|\bar{X})]PX_{ij}+u_{ij} \\ &= \alpha+\beta PX_{ij}+u_{ij} \end{aligned} \qquad (2.2)$$

where

$$\alpha = P(Y|\bar{X})$$

and

$$\beta = P(Y|X)-P(Y|\bar{X})$$

i.e., eq. (2.2) is a simple linear probability model. From this model it follows that

$$PY_{i.} = \alpha+\beta PX_{i.}+u_{i.} \qquad (2.3)$$

where

$$PX_{i.} = \frac{1}{N_i}\sum_{j=1}^{N_i} PX_{ij} \; ; \quad i=1,2,\ldots,n$$

i.e., the average probability for choosing party X in district No. i at election 1. Similarly $PY_{i.}$ is the average probability for choosing party Y at election 2 and $u_{i.}$ is the average stochastic disturbance in district No. i.

Since the number of voters in each district usually is so large that very little "sampling error" is attached to the proportion of voters PX_i voting for party X and the proportion PY_i voting for party Y, these proportions can safely be substituted for the corresponding average probabilities $PX_{i.}$ and $PY_{i.}$. Hence

$$PY_i = \alpha+\beta PX_i+u_i \qquad (2.4)$$

where u_i is nearly identical to $u_{i.}$.

Since α and β can be estimated from ecological data by ordinary regression methods, it is also possible to estimate the conditional probabilities in eq. (2.2) by

$$\hat{P}(Y|X) = \hat{\alpha}+\hat{\beta}$$
$$\hat{P}(Y|\bar{X}) = \hat{\alpha} \qquad (2.5)$$

as indicated in figure 2.1.

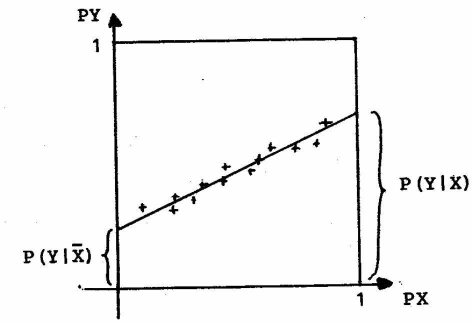


Fig. 2.1. Ecological estimation of individual voting behaviour P(Y|X) and P(Y|X̄)

Furthermore, since

$$\hat{P}(\bar{Y}|X) = 1-\hat{P}(Y|X) = 1-\hat{\alpha}-\hat{\beta}$$
$$\hat{P}(\bar{Y}|\bar{X}) = 1-\hat{P}(Y|X) = 1-\hat{\alpha} \qquad (2.6)$$

the proportions of all four possible combinations of voter mobility between the parties from election 1 to election 2 can be estimated from the results obtained by the ecological regression analysis.

Unfortunately, although the methodology is beautifully simple, the results obtained by the methodology is often far from adequate.

For instance, consider the empirical example in section 1. According to figure 1.1 there is a positive covariation on the ecological level between the proportion PX voting for the Social Democratic Party in 1971 and the proportion PY voting for the same party in 1973. With ordinary regression methods* the estimated intersection $\hat{\alpha}$ with the PY-axis is -.05 while the estimated slope $\hat{\beta}$ is .83. From eqs. (2.5) follows that the proportion of former Social Democrats voting for the same party again in 1973 is 78 pct. while the proportion of former non Social Democrats voting for the Social Democratic party in 1975 is -5 pct. Further, according to eqs. (2.6), the proportion of former Social Democrats leaving the party is 22 pct., while the proportion of stable non-Social Democrats is 105 pct.! These are clearly absurd results since proportions cannot be greater than 100 pct. or less than 0 pct.

This example is far from atypical. For another example consider the ecological covariation in chapter 2, figure 1.5, which also according to the simple linear regression model indicates that a negative number of voters moved from the Labour Party to the Conservative Party at the British General Elections 1964-66. Similar results are found in Hammond, Crewe and Payne, Miller, (op.cit.); Taylor and Johnson (1979, pp. 87-88).

In the ecological inference litterature quoted in section 1 there is general consensus that failure of the simple regression methodology stems from the fact that the individual-level model (2.2) is misspecified, but not from inefficiency of the regression approach as such. Hence much effort has been put on refining the regression approach.

2.2 Ecological Inference by more Complex Regression Models

It is a general finding in many ecological studies that it is the presence of "structural (or contextual) effects" (cf. Lincoln and Zeitz, 1980) which explains the poor performance of the simple regression model. In the simple regression model, eq. (2.1) it is naively assumed that the conditional probabilities $P(Y|X)$ and $P(Y|\bar{X})$ are constant across all individuals in all districts or at least randomly distributed around constant mean values. But all empirical evidence suggests that in most cases they are systematically related to structural (or contextual) variables characterizing the ecological units.

To overcome the problem of structural effects, two strategies have been developed (cf. Goodman, 1959; Stokes, 1969; Shively, 1969). One strategy is to divide the ecological units into separate homogeneous regions and then estimate individual behaviour inside each region with the simple regression model. The other strategy is to formulate multiple regression models which include the structural effects.

The first strategy is usually abandoned because the number of ecological units in each homogenous region tends to be too small to secure stable results. This, however, suggests that the number of units and hence the size of the single unit is an important matter in ecological research.

The other strategy has been more prominent, especially in electoral ecological research. Usually the multiple regression modelling takes its point of departure in the simple regression model, eq. (2.4), at the ecological level:

$$PY_i = P(Y|\bar{X})+[P(Y|X)-P(Y|\bar{X})]PX_i+u_i \qquad (2.7)$$

Now, to account for structural effects, it is assumed that the conditional probabilities $P(Y|X)$ and $P(Y|\bar{X})$ vary across units, i.e.

$$PY_i = P_i(Y|\bar{X})+[P_i(Y|X)-P_i(Y|\bar{X})]PX_i+u_i \qquad (2.8)$$

A simple approach to eq. (2.8) is based on the assumption that the conditional probabilities are linearly related to PX_i:

$$P_i(Y|X) = a_o+a_1PX_i$$
$$P_i(Y|\bar{X}) = b_o+b_1PX_i \qquad (2.9)$$

Inserting (2.9) into (2.8) the model becomes:

$$PY_i = b_o+(a_o+b_o+b_1)PX_i+(a_1-b_1)PX_i^2+u_i \qquad (2.10)$$

i.e. the parabolic regression model. Unfortunately with this model it is not possible to estimate all four parameters in (2.9) since there are only three regression coefficients in (2.10). Employing data from Belgian elections Boudon (1963) introduced the hypothesis $b_1 = 0$ but he obtained unrealistic estimates of individual voting behaviour. Likewise Miller (1977, pp. 42-48), with data from British elections, discarded the possibility of ecological inference by the hypothesis $a_1 = b_1$; and instead he used the model to measure the amount of structural effect by comparing the ecological regression results with survey results.

Another approach to the problem of structural effects is to assume that the conditional probabilities are linearly related to a number of external variables (e.g. Crewe and Payne, 1976):

$$P_i(Y|\bar{X}) = a_o+a_1Q_{1i}+a_2Q_{2i}+\ldots+a_mQ_{mi}$$
$$P_i(Y|X)-P_i(Y|\bar{X}) = b_o+b_1R_{1i}+b_2R_{2i}+\ldots+b_nR_{ni} \qquad (2.11)$$

then, according to (2.8):

$$PY_i = a_o+a_1Q_{1i}+\ldots+a_mQ_{mi}$$
$$+b_oPX_i+b_1R_{1i}PX_i+\ldots+b_nR_{ni}PX_i+u_i \qquad (2.12)$$

i.e. the linear regression of PY_i on a set of variables $Q_1,\ldots, Q_m$ and the interaction between a set of variables $R_1,\ldots, R_m$ and PX. With this model Crewe and Payne succeeded in obtaining realistic estimates of the British two-party vote within two classes, manual and non-manual. However, with 10 independent variables in the regression equation together with 4 interaction terms one cannot help feeling that this result is not very stable. Furthermore the validity of the methodology deteriorated when the third party of "non-voting" was included in the analysis.

A third approach to overcome the problem of structural effects is called "constrained regression". It has been used for ecological estimation of individual voting behaviour in multi-party systems (Lewin et al. 1972; Hoschka and Schunck 1975). Similarly to the simple regression model (2.7) for two-party voting behaviour the initial model for multi-party voting behaviour disregards the possibility of structural effects:

$$PY_{ji} = P(Y_j|X_1)PX_{1i}+\ldots+P(Y_j|X_m)PX_{mi}+u_i; \quad j=1,2,\ldots,n \qquad (2.13)$$

where $P(Y_j|X_h)$ is the conditional probability for choosing party no. j at election no. 2, given that party no. h was chosen at election no. 1. Since the initial estimates of the conditional probabilities yield highly unrealistic results, the initial estimates are "forced" to more realistic results by introducing a number of constraints such as

$$0 \le P(Y_j|X_h) \le 1$$

and

$$\sum_{j=1}^{n} P(Y_j|X_h) = 1$$

and at the same time optimizing the fit of the model to the ecological data.

Using data from West German Bundestag-elections Hoschka and Schunck found a distressingly poor accord between the ecological estimates of individual voting behaviour and survey results, and they rejected the validity of the methodology. Lewin et. al. were more enthusiastic about their own findings on Swedish elections, but other researchers (Quessel 1973; Berglund 1974) have convincingly raised serious doubts about the validity of the methodology.

2.3 Introduction to the Latent Structure Approach

The conclusion from the brief survey above about the experience with today's most prominent approach for electoral ecological inference, the regression approach, is not very optimistic. In fact, mainly because of contamination from structural effects it seems nearly impossible to estimate individual voting behaviour from ecological data. At least this seems to be the case if the estimation concerns more than two parties or more than two social classes.

The strategy of dividing the ecological units into separate homogenous regions and thus avoiding structural effects may be wise, but this strategy requires data from many small ecological units. This, however, stresses the importance of the quality of the data collected for the purpose of ecological inference.

The regression approach has two very serious pitfalls always troubling ecological analysts. One is called the problem of mis-specification, the other the problem of multi-collinearity. The problem of mis-specification means that PY is not explained sufficiently by PX. This is usually solved by introducing more manifest variables into the regression equation. However, this procedure may lead directly to the other pitfall of multi-collinearity, meaning that the estimation of coefficients is very unstable if there are too many (internally correlated) regressors or independent variables (cf. Langbein and Lichtman, 1978).

In my opinion the dangerous journey between the Scylla of mis-specification and the Charybdis of multi-collinearity is caused by the fundamental asymmetric conception of PX and PY inherent in the regression approach, as if PX was the cause for PY. Instead I prefer the perfect symmetric approach inherent in latent structure analysis, very much akin to factor-analysis, where both PX and PY are considered two different functions of the same latent variables.

With this approach the explanation of the covariation between PX and PY is that the two variables have "something in common", i.e. the latent factors, as schematically displayed in figure 2.2.

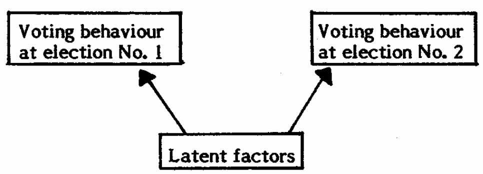


Figure 2.2 Schematic Presentation of the Latent Structure Approach

One example could be that the ecological covariation between proportion PX voting for the Social Democrats in 1971 and the proportion PY voting for the Social Democrats in 1973 in figure 1.1 is explained by the fact that a number of (unknown or latent) causes for Social Democratic voting are working both in 1971 and 1973, in spite of the bad luck in 1973. Not that PX is the cause for PY, possibly in connection with structural efects, as schematically displayed in figure 2.3.[1]

In the following sections a mathematical latent structure theory is proposed which shows that it is not always necessary to estimate the latent variables for making ecological inference. The important thing is to establish the functional relation between the individual and the ecological correlation.

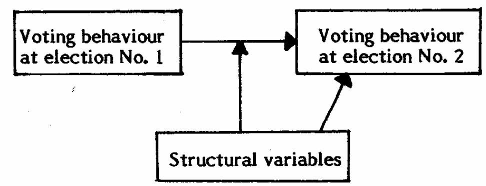


Figure 2.3 Schematic Presentation of the Regression Approach

At this point, readers familiar with the ecological inference litterature may wonder. As mentioned in section 1, if there is anything that the advocates as well as the critics of the practice of ecological inference agree upon it is that correlations are certainly not paving the golden road to cross-level inference.

For instance, take the classic example discussed by the harsh critic Robinson (1950, pp. 352-354). He compared the individual Pearson correlation between the two binary variables X: black-white and Y: illiterate-literate with the ecological Pearson correlation across US states between two corresponding average variables PX: proportion of blacks and PY: proportion of illiterates. The computed individual correlation was: .203, while the ecological correlation was: .773, so the divergence between the two correlation coefficients were considerable.

Working with the problem of ecological inference I often wondered why not one of the skilled scientists elaborating on this example (e.g. Hannan, 1971, pp. 63-66; Hammond, 1973, p. 77; Firebaugh, 1978, pp. 567-569; Hannushek et al., 1974) ever criticized the use of the Pearson correlation on binary data. As virtually every introductory textbook on applied statistics notes, the assumptions behind the product-moment Pearson correlation coefficient between two variables are that they are continuous interval scale variables and that the joint distribution of the variables is the two-dimensional normal distribution, which is certainly not the case for the two mentioned individual-level binary variables X and Y. For this situation many textbooks (e.g. Conolly and Sluckin, 1971, pp. 174-178) propose the socalled tetrachoric (or fourfold, or 2 x 2 contingency table) correlation which assumes that each binary variable is produced by cutting a latent normal distributed continuous variable in two parts. The tetrachoric correlation is then simply the product-moment correlation between these two latent variables.

For the proportions PX and PY on the ecological level the assumptions behind the Pearson correlation are often more realistic, but with very high or very low percentages, the logit correlation, i.e. the Pearson correlation between logit-transformed percentages, is suggested as a better choice (cf. section 3.4).

It is amusing to discover that Robinson's example denies the ecological fallacy. Using Robinson's data it actually turns out that tetrachoric correlation (computed with formula (3.45) in section 3.5) is .747, which is only slightly below the ecological Pearson correlation which, as mentioned, is .773. However, in section 3.5 I shall discuss why this result is rather a joke than a scientific result.

Returning to more serious business in the next sections it is within the framework of latent structure theory specified under what conditions we should expect that the individual tetrachoric correlation is approximately equal to the ecological logit correlation.

3 The Latent Structure Theory for Binary Choice

Latent structure theory was founded by Lazarsfeld (1950, 1954; Lazarsfeld and Henry, 1968) and developed in psychometrics through several decades (Andersen, 1980b). It has been applied to individual voting behaviour (Wiggins, 1973), but to my knowledge it has not prior to the work presented in this study been applied to ecological inference.

At first a latent structure model for individual voting behaviour is formulated and deductions are made from the individual to the aggregate level. These deductions can in turn be applied to ecological inference.

Strictly speaking the theory applies to genuine binary choice only, that is, the voter has the choice between two and only two categories at each election. This situation is not very realistic, because even in two-party systems the third alternative of "non-voting" is a possible choice. The methodology for binary choice can, however, as demonstrated in section 4.1, with some approximation be applied to analyse "crude binary choice" which means that one party or group of parties are considered as one alternative and the "other" parties (incl. the "party" of non-voting) are considered as the second alternative. Later on in section 4.3, the methodology is generalized to multiple choice by simultaneously considering multiple subsets of binary choice.

3.1 Individual Binary Choice

Latent structure theory assumes that the probability of a certain choice is a function of a latent variable associated with each individual. With regard to voting behaviour the present model assumes that the probability of the voters' choice of party X at election No. 1 and the choice of party Y at election No. 2 are functions of the same value on the latent variable θ at both elections:[1]

$$P(X|\theta) = f(\theta) \qquad P(Y|\theta) = g(\theta) \tag{3.1}$$

Furthermore the fundamental axiom of "local independence" in latent structure theory states:

$$P(X\cap Y|\theta) = P(X|\theta)P(Y|\theta) \tag{3.2}$$

that is, the two responses X and Y are statistically independent, given the latent trait θ[2].

Originally the probability model selected for individual behaviour was the Linear Logit Model, which in chapter 2 served so well for ecological analysis. On the individual level the probabilities are

$$P(X|\theta) = \Psi(\alpha_o+\alpha\theta) \qquad P(Y|\theta) = \Psi(\beta_o+\beta\theta) \tag{3.3}$$

where Ψ is the logistic function

$$\Psi(t) = \frac{\exp(t)}{1+\exp(t)} \tag{3.4}$$

For the one-dimensional case, θ is interpreted as the voters "latent opinion" or position on a latent political cleavage-dimension. At election 1 the parameter α associated with the party X is the position of the party on this same dimension. The other party-parameter α_o, which does not interact with the voters' latent opinion, is interpreted as the "general popularity" of party X independent of political cleavages. Similarly at election 2 β is the position of party Y on the latent cleavage dimension and β_o is the party's "general popularity".

In contrast to psychometric test situations the assumption of one-dimensionality is not very realistic for political voting behaviour where several cleavage-dimensions can influence the voters' choice. For this reason the model (3.3) is generalized to the multi-dimensional case by interpreting θ as a vector of variables:

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_k \end{pmatrix}$$

where θ_1 is the voters' position on latent cleavage dimension No. 1, θ_2 is the position on cleavage dimension No. 2, etc. for all k dimensions. Similarly α for party X at election No. 1 and β for party Y at election No. 2 are vectors of parameters indicating the party's positions:

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$$

$$\beta = (\beta_1, \beta_2, \ldots, \beta_k)$$

on all k cleavage dimensions. Furthermore, the parameter products $\alpha\theta$ and $\beta\theta$ are the vector products:

$$\alpha\theta = \alpha_1\theta_1+\alpha_2\theta_2+ \ldots+\alpha_k\theta_k$$

$$\beta\theta = \beta_1\theta_1+\beta_2\theta_2+ \ldots+\beta_k\theta_k$$

But still, in the multi-dimensional case α_o and β_o in eq. (3.3) are one-dimensional parameters for general popularity.

Unfortunately, it turned out that the mathematical structure of the logistic function is not well suited for aggregate mathematical derivation, especially not in the multidimensional case (cf. Gumbel, 1961). For this technical reason another function, which is a very close approximation to the logistic function, is applied. It is the well known standard normal cumulative distribution function:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp(-\tfrac{1}{2}s^2)\,ds \qquad (3.5)$$

It has been shown (Lord and Novick, 1968, p. 399) that by dividing the argument in the function Φ with the constant 1.7, the absolute difference between Φ and Ψ is always less than 1 percentage point:

$$|\Phi(t/1.7)-\Psi(t)| < 0.01 \quad \text{for} \quad -\infty < t < +\infty \qquad (3.6)$$

So for all practical means, apart from rescaling of the argument the results obtained by the two different functions cannot be distinguished. This means that the following probability model:

$$P(X|\theta) = \Phi(\alpha_o+\alpha\theta)$$
$$P(Y|\theta) = \Phi(\beta_o+\beta\theta) \qquad (3.7)$$

apart from rescaling is a close approximation to the model (3.3). The model (3.7) is called "the normal ogive model" in psychometrics (Lord and Novick, 1968, p. 366), and "the linear probit model" in econometrics. As implied above the probit model is only used as a "stand-in" for the logit model for technical reasons. Later on, when the methodology is generalized to multiple choice in section 4, the logit model will be reintroduced.

The parameters of the model (3.3) (or (3.7)) are undetermined unless certain constraints are introduced, e.g.

$$\sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \beta_i = 0 \qquad (3.8)$$

The constraints on θ are introduced in the next section, which states the assumptions about the population distribution of θ both within and between election districts.

3.2 Population Distribution in the Latent Space

The models introduced in the last section assume functional homogeneity, which means that the party parameters α_o, α, β_o, β are the same for all individuals. This assumption is not realistic for the country taken as a whole, because of different "political culture" in different parts of the country. For this reason the model is only assumed to be valid within relatively politically homogeneous geographical regions. Consequently, the perequisite for making ecological inference is that each single political region can be further subdivided into local districts, as schematically illustrated in figure 3.1.

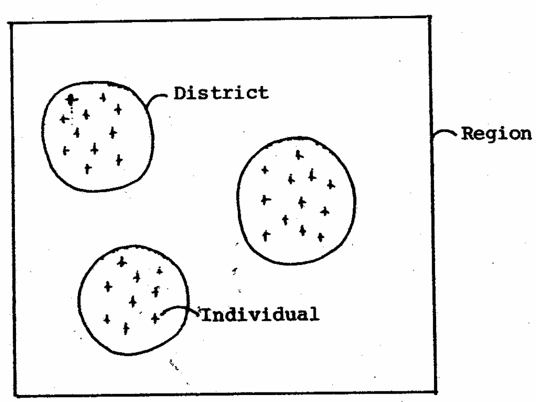


Figure 3.1. Subdivision of a politically homogeneous region

Within each district it is assumed that the individual population distribution of θ in the latent space follows the k-dimensional normal distribution:

$$\theta \sim N_k(\mu, \sigma) \tag{3.9}$$

with the vector of mean values:

$$\mu \doteq \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}$$

where

$$\mu_i = E(\theta_i)$$

and with the covariance matrix:

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

where

$$\sigma_{ij} = \mathrm{Cov}(\theta_i, \theta_j) \quad ; \quad i \neq j$$

and

$$\sigma_{ii} = \mathrm{Var}(\theta_i) \quad ; \quad i = 1, 2, \ldots, k$$

Across districts within the region it is arbitrarily assumed that the vector of mean values μ follows the k-dimensional standard normal distribution:

$$\mu \sim N_k(\underline{0}, I) \tag{3.10}$$

with the vector of mean values:

$$\underline{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and the covariance matrix:

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

The constraint (3.10) on the vector of mean values is the same as arbitrarily introduced for the ecological linear logit in chapter 2, eq. (2.5). Consequently the covariance-matrix σ describes the relation between the within-district individual variation and the between-district ecological variation in the latent space.

It is the assumptions about the properties of σ, i.e. the assumptions about the relation between individual and ecological variation, which puts the final constraints on the parameters of the model (3.3). It follows from the within-district distribution assumption (3.9) and the across-district distribution assumption (3.10) that the total within-region individual distribution of θ is

$$\theta \sim N_k(\underline{0}, I+\sigma) \qquad (3.11)$$

i.e. the k-dimensional normal distribution with the vector of mean values $\underline{0}$ and the covarance matrix $I+\sigma$.

I have experimented with many more or less complex assumptions about σ (Thomsen, 1975) and finally arrived at the conclusion, that a very satisfactory solution to the problem of ecological inference can be obtained with the most simple of all possible solutions, which I call "the assumption of isomorphism". The assumption of isomorphism is

$$\sigma = \kappa I \qquad (3.12)$$

which means that the ratio between the within-districts individual variance and the between districts ecological variance is equal to the constant κ on all k dimensions in the latent space:[3]

$$\kappa = \frac{Var(\theta_1)}{Var(\mu_1)} = \frac{Var(\theta_2)}{Var(\mu_2)} = \dots = \frac{Var(\theta_k)}{Var(\mu_k)}$$

Consequently in the special case of isomorphism, the properties of σ is determined by one single value: the value of the variance ratio κ.

A verbal interpretation of isomorphism is that the variation between individuals has the same structure as the variation between districts. The concept is borrowed from crystallography where isomorphism means "similarity of structure, especially between the crystals of different chemical substances" (Foreman, 1959, p. 271).

In the special case of isomorphism, the within-districts distribution of θ is

$$\theta \sim N_k(\mu, \kappa I) \qquad (3.13)$$

and the total within-region distribution of θ is

$$\theta \sim N_k(\underline{0}, (1+\kappa) I) \qquad (3.14)$$

In the following section the model for aggregate behaviour is mathematically derived from the distribution assumptions about θ as well for the general case of any σ as for the special case of isomorphism.

3.3 Mathematical Derivations[4]

To carry through the mathematical derivations, we need to know the distribution of $u = \alpha\theta$ and $v = \beta\theta$ (cf. Anderson, 1958). It follows from the distribution assumption (3.9) that the joint distribution of u and v within each district is

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \alpha\theta \\ \beta\theta \end{pmatrix} \sim N_2\left[\begin{pmatrix} \alpha\mu \\ \beta\mu \end{pmatrix}, \begin{pmatrix} \alpha\sigma\alpha' & \alpha\sigma\beta' \\ \alpha\sigma\beta' & \beta\sigma\beta' \end{pmatrix}\right] \qquad (3.15)$$

i.e., the two-dimensional normal distribution with mean vector and covariance-matrix as indicated. α' is the transposed α-vector, and β' is the transposed β-vector.

In the following derivations it is assumed that the number of voters in each district is so large that very little "sampling" error is attached to the proportion of voters voting for each party in each district. Consequently, we assume that the proportion choosing party X at election 1 in a district is equal to the mean individual probability for choosing the party:

$$PX = P(X|\mu,\sigma) = \int_\theta P(X|\theta)\,p(\theta|\mu,\sigma)\,d\theta$$

$$= \int_u P(X|u)\,p(u|\mu,\sigma)\,du$$

$$= \int_u \Phi(\alpha_o+u)\frac{1}{\sqrt{\alpha\sigma\alpha'}}\varphi\left(\frac{u-\alpha\mu}{\sqrt{\alpha\sigma\alpha'}}\right)du$$

$$= \Phi\left(\frac{\alpha_o+\alpha\mu}{\sqrt{1+\alpha\sigma\alpha'}}\right) \qquad (3.16)$$

where φ is the standard normal density distribution function. Equation (3.16) is derived by weighting the model (3.7) with the distribution assumption (3.9) (cf. Lord and Norvick, 1968, pp. 376-77). Analogous to eq. (3.16) the proportion choosing party Y at election 2 is

$$PY = P(Y|\mu,\sigma) = \Phi\left(\frac{\beta_o+\beta\mu}{\sqrt{1+\beta\sigma\beta'}}\right) \qquad (3.17)$$

Because α,β,σ are assumed constant across districts, these results show that the model (3.16) and (3.17) for aggregate behaviour apart from rescaling of the arguments with the constants $1/\sqrt{1+\alpha\sigma\alpha'}$ and $1/\sqrt{1+\beta\sigma\beta'}$ is equal to the model (3.7) for individual behaviour.

We shall now establish the dependence of individual voting behaviour between election 1 and election 2 by application of the axiom of local independence, eq. (3.2). With this axiom the dependence between X and Y is explained by the variation in the latent trait θ.

The joint proportion PXY of voters voting for party X at election 1 and for party Y at election 2 in a district is found by the mean joint probability across individuals (c.f. Bartholomew, 1980, p. 310):

$$PXY = P(X\cap Y|\mu,\sigma)$$

$$= \int_\theta P(X|\theta)P(Y|\theta)\,p(\theta|\mu,\sigma)\,d\theta$$

$$= \int_u\int_v \Phi(\alpha_o+u)\Phi(\beta_o+v)\,p(u,v|\mu,\sigma)\,\partial u\partial v$$

$$= \Phi_2(x,y,\rho_w) \qquad (3.18)$$

where Φ_2 is the two-dimensional standard normal cumulative distribution function with the arguments:

$$x = \Phi^{-1}(PX) = \frac{\alpha_o+\alpha\mu}{\sqrt{1+\alpha\sigma\alpha'}} \qquad (3.19)$$

and

$$y = \Phi^{-1}(PY) = \frac{\beta_o+\beta\mu}{\sqrt{1+\beta\sigma\beta'}} \qquad (3.20)$$

known as probit-transformations of PX and PY, and with the latent within-district correlation coefficient:

$$\rho_w = \frac{\alpha\sigma\beta'}{\sqrt{1+\alpha\sigma\alpha'}\sqrt{1+\beta\sigma\beta'}} \qquad (3.21)$$

known as "the tetrachoric correlation coefficient" (Johnson and Kotz, 1973, pp. 117-20).

In the special case of isomorphism, cf. equation (3.12), the latent within-districts correlation coefficient becomes:

$$\rho_w = \frac{\kappa\alpha\beta'}{\sqrt{1+\kappa\alpha\alpha'}\sqrt{1+\kappa\beta\beta'}} \qquad (3.22)$$

Similarly to eq. (3.18) the total joint proportion PXY_o of voters in all districts inside the region voting for party X at election 1 and for party Y at election 2 is found by the distribution assumption (3.11):

$$PXY_o = \Phi_2(x_o, y_o, \rho_t) \tag{3.23}$$

where

$$x_o = \Phi^{-1}(PX_o) = \frac{\alpha_o}{\sqrt{1+\alpha(I+\sigma)\alpha'}} \tag{3.24}$$

$$y_o = \Phi^{-1}(PY_o) = \frac{\beta_o}{\sqrt{1+\beta(I+\sigma)\beta'}} \tag{3.25}$$

$$\rho_t = \frac{\alpha(I+\sigma)\beta'}{\sqrt{1+\alpha(I+\sigma)\alpha'}\sqrt{1+\beta(I+\sigma)\beta'}} \tag{3.26}$$

PX_o and PY_o are the marginal proportions voting for X at election 1 and for Y at election 2 in the whole region, and ρ_t is the total latent correlation coefficient.

In the special case of isomorphism the total latent correlation coefficient becomes

$$\rho_t = \frac{(1+\kappa)\alpha\beta'}{\sqrt{1+(1+\kappa)\alpha\alpha'}\sqrt{1+(1+\kappa)\beta\beta'}} \tag{3.27}$$

In equation (3.23) x_o and y_o can be computed from the proportions PX_o and PY_o, taken from official election statistics applying eqs. (3.24) and (3.25). So the only information needed to estimate the proportion PXY_o of voters moving from party X to party Y from election 1 to election 2 is the total latent correlation coefficient ρ_t.

Section 3.5 discusses under what circumstances ρ_t can be inferred from ecological data.

3.4 Ecological Analysis

We shall now reenter the realm of ecological analysis. According to (3.19) and (3.20) the probit to the proportions PX and PY in district No. i is

$$x_i = \Phi^{-1}(PX_i) = \frac{\alpha_o+\alpha\mu_i}{\sqrt{1+\alpha\sigma\alpha'}} \tag{3.28}$$

$$y_i = \Phi^{-1}(PY_i) = \frac{\beta_o+\beta\mu_i}{\sqrt{1+\beta\sigma\beta'}} \tag{3.29}$$

Because Φ except for a scale constant is a very close approximation to Ψ (cf. eq. (3.6)) it follows that the model (3.28) is a close approximation to the ecological linear logit model in chapter 2, eq. (2.5). In fact the model (3.28) above can be transformed to the model (2.5) in chapter 2 by the approximation (3.6) and the transformations:

$$LX_i = x_i/1.7$$

$$a_o = \alpha_o/(1.7\sqrt{1+\alpha\sigma\alpha'}) \tag{3.30}$$

$$a_j = \alpha_j/(1.7\sqrt{1+\beta\sigma\beta'}) \; ; \; j = 1,2, \ldots, k$$

As discussed in chapter 2, section 2.1, the stochastic component e_i in the model (2.5) in chapter 2 can be disregarded within politically homogenous regions.

Disregarding "sampling errors" the distribution assumption (3.10) together with (3.28) and (3.29) leads to the following expectations and variances across districts for x and y:

$$E(x) = x_. = \frac{1}{N}\sum_{i=1}^{n} N_i x_i = \frac{\alpha_o}{\sqrt{1+\alpha\sigma\alpha'}} \tag{3.31}$$

$$E(y) = y_. = \frac{1}{N}\sum_{i=1}^{n} N_i y_i = \frac{\beta_o}{\sqrt{1+\beta\sigma\beta'}} \tag{3.32}$$

$$Var(x) = \frac{1}{N}\sum_{i=1}^{n} N_i (x_i - x_.)^2 = \frac{\alpha\alpha'}{1+\alpha\sigma\alpha'} \tag{3.33}$$

$$Var(y) = \frac{1}{N}\sum_{i=1}^{n} N_i (y_i - y_.)^2 = \frac{\beta\beta'}{1+\beta\sigma\beta'} \tag{3.34}$$

where n is the number of districts and N is the number of voters in the region. N_i is the number of voters in district No. i.

From the distribution assumption (3.10) it is also possible to find the covariance across districts between x and y:

$$Cov(x,y) = \frac{1}{N}\sum_{i=1}^{n} N_i (x_i - x.)(y_i - y.)$$

$$= \frac{\alpha\beta'}{\sqrt{1+\alpha\sigma\alpha'}\sqrt{1+\beta\sigma\beta'}} \quad (3.35)$$

and the ecological correlation coefficient:

$$\rho_e = \rho(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)}\sqrt{Var(y)}}$$

$$= \frac{\alpha\beta'}{\sqrt{\alpha\alpha'}\sqrt{\beta\beta'}} \quad (3.36)$$

The ecological probit correlation (3.36) is computed from election statistics by the Pearson correlation across districts between probit-transformed proportions of voters voting for party X at election 1 and the probit-transformed proportion of voters voting for party Y at election 2:

$$r_e = r_e(x,y) = \frac{\Sigma N_i (x_i - x.)(y_i - y.)}{\sqrt{\Sigma N_i (x_i - x.)^2}\sqrt{\Sigma N_i (y_i - y.)^2}} \quad (3.37)$$

Because logit-transformed proportions (cf. eq. (3.6)) except for a scale constant, are very close to probits, we might as well instead of probits use logits, which are easier to compute, by computation of the ecological correlation (3.37). We might even instead of x and y use the proportions PX and PY by computation of the ecological correlation because they are approximately linearly related to probits if the range is not too close to either 0 or 1.

It is an important feature of the theory that the ecological correlation in the one-dimensional case according to eq. (3.36) is either +1 or -1. This means that

ecological correlations less strong than ±1 are explained not by any kind of "error-component" but by multi-dimensionality as if it was possible to explain all ecological variations by taking just enough dimensions into account.

As discussed in chapter 2, the ecological logit correlations have been used for factor analysis of election results for each party at Danish parliamentary elections since 1920. The two most important cleavage dimensions in the latent space were identified as the rural-urban dimension and the middle class-working class dimension respectively (Thomsen, 1979).

3.5 Ecological Inference

According to eqs. (3.23)-(3.26) the proportion PXY_o of all voters in the region moving from party X to party Y between two elections can be regarded as a function F of the marginal proportions PX_o and PY_o and the total latent correlation-coefficient ρ_t:

$$PXY_o = \Phi_2(x_o, y_o, \rho_t)$$

$$= F(PX_o, PY_o, \rho_t) \quad (3.38)$$

Table 3.1. shows the notation for the four different proportions of voter mobility from voting either X or $\bar{X}$ (non-X) at election 1 to voting either Y or $\bar{Y}$ (non-Y) at election 2.

Table 3.1 Binary Choice Table: Notation for Proportions of All Voters

Choice at Election 1	Choice at Election 2		
	Y	$\bar{Y}$ (non-Y)	Total
X	PXY_o	$PX\bar{Y}_o$	PX_o
$\bar{X}$ (non-X)	$P\bar{X}Y_o$	$P\bar{X}\bar{Y}_o$	$P\bar{X}_o$
Total	PY_o	$P\bar{Y}_o$	1

Since we in equation (3.38) know the function F and the marginal proportions PX_o and PY_o, the only information needed to estimate the content of table 3.1. is ρ_t.

According to eq. (3.26) ρ_t is a function of α, β and σ. If we know the covariance-matrix σ; i.e. the relation between the individual and the ecological variation in the latent space, then the party-vectors α and β can be estimated by factor-analytical methods. Unfortunately, it is not possible to estimate σ from ecological data, which is a conclusion similar to Robinson's "ecological fallacy", discussed in section 1.

If we introduce the hypothesis of isomorphism, eq. (3.12), then the problem of ecological inference is boiled down to finding the variance ratio κ. Working with the assumption of isomorphism, I tried to estimate κ by comparing ecological correlations with individual tetrachoric correlations from survey results, but I suddenly realized that it is not necessary to estimate the variance ratio κ if only it is sufficiently high! This can be seen by rewriting eqs. (3.22) and (3.27):

$$\rho_w = \frac{\alpha\beta'}{\sqrt{1/\kappa+\alpha\alpha'}\sqrt{1/\kappa+\beta\beta'}} \qquad (3.39)$$

$$\rho_t = \frac{\alpha\beta'}{\sqrt{1/(1+\kappa)+\alpha\alpha'}\sqrt{1/(1+\kappa)+\beta\beta'}} \qquad (3.40)$$

If we compare eq. (3.39) for the within-district correlation ρ_w, and eq. (3.40) for the total individual correlation ρ_t with eq. (3.36) for the ecological correlation ρ_e, the relative strength is

$$|\rho_w| < |\rho_t| < |\rho_e| \qquad (3.41)$$

that is, given the hypothesis of isomorphism, the ecological correlation is always stronger than the total individual correlation which again is always stronger than the within-districts correlation (if they are not all equal to 0).

This result indicates that by using ecological correlation as substitute for individual tetrachoric correlation we estimate a dependence between X and Y, which is at least as strong as the true dependence. Furthermore, when the variance

ratio is very high the individual correlations converge towards the ecological correlation:[5]

$$\lim \rho_w = \lim \rho_t = \rho_e \quad \text{for} \quad \kappa \to \infty \qquad (3.42)$$

A high variance ratio κ means that the within-districts variation between individuals is considerably higher than the ecological variation between districts; and this creates the convergence between the within-districts and the total correlation. In Robinson's opinion this situation is most unlikely because "... all available evidence is that (whatever properties X and Y may denote) the correlation between X and Y is certainly not larger for relative homogeneous sub-groups of persons than it is for the population at large." (Robinson, 1950, pp. 356). In my opinion this is not always the case. At least for electoral behaviour within homogeneous political regions the within-districts correlation often approximates the total individual correlation. This is for example illustrated at Danish elections where surveys made by the Danish Television Company in a single district within a few political regions give valid estimates of national voter mobility and class voting. That indicates that with regard to voter mobility and class voting each district is a miniature picture of the whole political region.

Consequently, the hypothesis behind the ecological estimation of individual electoral behaviour is

$$\rho_w = \rho_t = \rho_e \qquad (3.43)$$

that is, the ecological correlation can be substituted for the individual tetrachoric correlation in formula (3.38):

$$PXY_o = F(PX_o, PY_o, \rho_e) \qquad (3.44)$$

A computer algorithm for the function F has been developed by Brown (1977) but it has turned out to be too time-consuming fo the iteration procedure described later on for multiple choice. In this case Yule's well-known approximation to the tetrachoric correlation (Johnson and Kotz, 1972, p. 119) is used. With the terminology from table 3.1 it is defined

$$r = \frac{(PXY_oP\bar{X}\bar{Y}_o) - (PX\bar{Y}_oP\bar{X}Y_o)}{(PXY_oP\bar{X}\bar{Y}_o) + (PX\bar{Y}_oP\bar{X}Y_o)} \qquad (3.45)$$

Solving this equation with respect to PXY_o (and using the notation from table 3.1) the solution is

$$PXY_o = \frac{1+2rPX_o+2rPY_o-r-\sqrt{(1+2rPX_o+2rPY_o-r)^2-8r(1+r)PX_oPY_o}}{4r}$$

$$\doteq G(PX_o, PY_o, r) \qquad (3.46)$$

With this formula, an approximate ecological estimate $\hat{PXY}_o$ of the individual behaviour PXY_o is obtained by inserting the ecological correlation r_e computed by eq. (3.37) into eq. (3.46):

$$PXY_o = G(PX_o, PY_o, r_e) \qquad (3.47)$$

Using eq. (3.47) on Robinson's classic example (1950, p. 353) of the ecological fallacy it is, as mentioned in section 2.3, amusing to discover that Robinson's example actually denies the ecological fallacy. In Robinson's example the proportion of blacks PX_o in 1930 was 9.55 pct and the proportion of illiterates PY_o was 4.03 pct. Inserting the ecological correlation across states $r_e = .773$ into eq. (3.47) the total proportion of all individuals which are both black and illiterate, is estimated 1.66 pct. while the true value is 1.55 pct., which is only slightly below the ecological estimate. Admittedly this is a joke rather than a scientific result. We really should not expect that ecological inference is possible with such large ecological units as states. But at least Robinson's example is not so well chosen because it does not prove the ecological fallacy. Ecological inference is more realistic for such small units as neighbourhoods or municipalities within homogeneous regions where the variation across areas has the same structure as the variation across individuals.

In the next section the estimation formula (3.47) for binary choice shall be applied in the context of multiple choice.

4 Ecological Inference For Multiple Choice

The greatest obstacle for using the latent structure theory for binary choice on Danish elections is that the voters' choice is certainly not genuine binary. For example at Danish parliamentary elections in the 1970s the number of parties having candidates in each district ranged from 11 to 13 (incl. the "party" of non-voting). In spite of that the theory for binary choice can be used in the context of multiple choice.

In subsection 4.1 the methodology for binary choice is applied to estimate "crude binary choice" which means that one party is considered as one alternative while the "other" parties are considered as the second alternative. Later on, after the introduction of the model for multiple choice in subsection 4.2 a more sophisticated methodology is developed in subsection 4.3 by simultaneously considering multiple subsets of genuine binary choice.

4.1 Estimation of Crude Binary Choice

Table 4.1 shows the notation for the m x n different proportions of voter mobility from voting for one of the m different parties (incl. the "party" of non-voting) at election 1 to voting for one of the n parties at election 2.

The problem of ecological inference is that we do not know these proportions - we only know the marginal proportions (the totals) in table 4.1 for each geographical district from official election statistics.

The methodology for binary choice has been used for "crude binary estimation" of voter mobility between Danish political parties. With the terminology in table 4.1 the joint proportion P_{hj} voting for party h at election 1 and for party j at election 2 is considered as a result of a crude binary choice between party X=h and all other parties at election 1 and a crude binary choice between party Y=j and all other parties at election 2. With formula (3.47) the proportion P_{hj} is estimated by

$$\hat{P}_{hj} = G(P_{h.}, P_{.j}, r_{hj}) \; ; \; h=1,2,\ldots,m \; ; \; j=1,2,\ldots,n \qquad (4.1)$$

where $P_{h.}$ is the proportion of all voters voting for party h at election 1 and $P_{.j}$ is the proportion of all voters voting for party j at election 2. r_{hj} is the ecological logit correlation between the proportion voting for party h at election 1 and the proportion voting for party j at election 2 across all election districts within a homogenous region.

<u>Table 4.1</u> Multiple Choice Table. Notation for Proportions of All Voters moving between the Parties from Election 1 to Election 2 within a homogenous Region

X: Choice at	Y: Choice at Election 2							
Election 1	1	2	...	j	...	n	Total
1	P_{11}	P_{12}	...	P_{1j}	...	P_{1n}	$P_{1.}$
2	P_{21}	P_{22}	...	P_{2j}	...	P_{2n}	$P_{2.}$
.							.
.	....						.
.							.
h	P_{h1}	P_{h2}	...	P_{hj}	...	P_{hn}	$P_{h.}$
.							.
.	....						.
.							.
m	P_{m1}	P_{m2}	...	P_{mj}	...	P_{mn}	$P_{m.}$
Total	$P_{.1}$	$P_{.2}$	...	$P_{.j}$	...	$P_{.n}$	1

Unfortunately this estimation procedure gives a tendency towards over-estimation of the mobility proportions so that

$$\sum_{j=1}^{n} \hat{P}_{hj} > P_{h.} \quad \text{and} \quad \sum_{h=1}^{m} \hat{P}_{hj} > P_{.j}$$

This bias is at least partly explained by the fact that the binary choice is crude and not a genuine choice between two homogeneous alternatives. The bias is mathematically explained in the next section by introducing the model for multiple choice.

4.2 The Model for Multiple Choice

The model for individual multiple choice, called the linear multiple logit model, can be regarded as a generalization of the model for individual binary choice (3.3) presented in subsection 3.1.

At election 1 each voter has one choice between m parties (incl. the "party" of non-voting). This choice is denoted by the stochastic variable X which has the possible values 1,2,...,m. Simplified the outcome X = h is written X_h. Analogous the choice of party j at election 2 is written Y_j.

Similarly to the binary choice model the probability for the outcome X_h as well as the probability for the outcome Y_j, are functions of the same latent vector of varibles θ, and the outcomes are assumed statistically independent, given the latent trait θ. On the individual level the probabilities are

$$P(X_h | \theta) = \exp(\alpha_{oh} + \alpha_h \theta)/C(\theta)$$
$$P(Y_j | \theta) = \exp(\beta_{oj} + \beta_j \theta)/D(\theta) \qquad (4.2)$$

where

$$C(\theta) = \sum_{h=1}^{m} \exp(\alpha_{oh} + \alpha_h \theta)$$
$$D(\theta) = \sum_{j=1}^{n} \exp(\beta_{oj} + \beta_j \theta) \qquad (4.3)$$

As for the binary model θ is interpreted as the voters latent position on k latent political cleavage dimensions, and

$$\alpha_h = (\alpha_{1h}, \alpha_{2h}, \ldots, \alpha_{kh}) \; ; \; h=1,2, \ldots, m$$

is the position of party h on these same dimensions at election 1, while

$$\beta_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{kj}) \; ; \; j=1,2, \ldots, n$$

is the position of party j on the dimensions at election 2. α_{oh} and β_{oj} are one-dimensional parameters for general popularity of the two parties h and j at election 1 and election 2, respectively.

The constants across parties C(θ) and D(θ) are introduced to insure that

$$\sum_{h=1}^{m} P(X_h|\theta) = \sum_{j=1}^{n} P(Y_j|\theta) = 1$$

The following text analyses the mathematical structure of the model at election 1. At election 1 the probability for choosing party h is rewritten

$$P(X_h|\theta) = \Psi[f_h(\theta)] \tag{4.4}$$

where Ψ is the logistic function (3.4) and

$$f_h(\theta) = \alpha_{oh}+\alpha_h\theta-\ln[C(\theta)-\exp(\alpha_{oh}+\alpha_h\theta)] \tag{4.5}$$

In the special (unrealistic) case with only two possible parties, i.e. m = 2, equation (4.5) becomes

$$f_1(\theta) = (\alpha_{o1}-\alpha_{o2})+(\alpha_1-\alpha_2)\theta$$
$$= \alpha_o+\alpha\theta \tag{4.6}$$

and equation (4.4) becomes

$$P(X_1|\theta) = \Psi(\alpha_o+\alpha\theta) \tag{4.7}$$

This result proves that the model (4.2) for multiple choice is equal to the model (3.3) for genuine binary choice if m = 2.

In the general (realistic) case where m > 2, $f_h(\theta)$ is not a linear function of θ as in the special binary case. This is proved by forming first and second order derivatives with regard to each variable in the vector θ. The first derivative with respect to the i'th variable is

$$\frac{\partial f_h(\theta)}{\partial\theta_i} = \alpha_{ih}-E(\alpha_{i\bar{h}}|\theta) \tag{4.8}$$

where

$$E(\alpha_{i\bar{h}}|\theta) = \sum_{g\neq h} \alpha_{ig}P(X_g|\theta)/[1-P(X_h|\theta)] \tag{4.9}$$

i.e. the probability-weighted mean value of α_{ig} for all other parties than party h. Formula (4.8) is clearly not constant across individuals if m > 2, hence the function (4.5) is not a linear function of θ.

The second derivative with respect to the i'th variable is

$$\frac{\partial^2 f_h(\theta)}{\partial\theta_i^2} = -\mathrm{Var}(\alpha_{i\bar{h}}|\theta) \tag{4.10}$$

where

$$\mathrm{Var}(\alpha_{i\bar{h}}|\theta) =$$

$$\sum_{g\neq h} [\alpha_{ig}-E(\alpha_{i\bar{h}}|\theta)]^2 P(X_g|\theta)/[1-P(X_h|\theta)] \tag{4.11}$$

i.e. the probability-weighted variance of α_{ig} for all other parties than party h. Since (4.10) is always less than 0 if m > 2, it indicates that the function (4.5) is "curbing downwards" if m > 2.

The methodology for "crude binary choice" in subsection 4.1 is based on the crude assumption that $f_h(\theta)$ can be approximated to a linear function of θ, although this is not realistic for m > 2. According to the distribution assumption (3.11) about the total individual distribution of θ and the first derivative (4.8) the best linear approximation to $f_h(\theta)$ through the origin of the latent space is

$$f_h(\theta) \approx f_h(\underline{0}) + \sum_{i=1}^{k} [\alpha_{ih} - E(\alpha_{i\bar{h}}|\underline{0})]\theta_i \qquad (4.12)$$

This approximation is only perfect if m = 2, as demonstrated in figure 4.1 Otherwise the function $f_h(\theta)$ is "curbing downwards", cf. eq. (4.10).

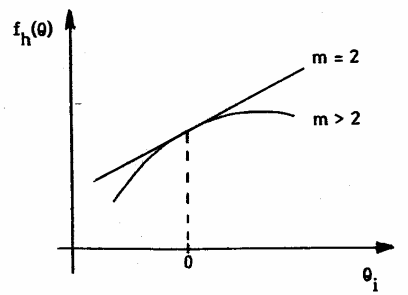


Figure 4.1. The relation between the function $f_h(\theta)$ and θ_i for m = 2 and m > 2 respectively.

Since P $(X_h|\theta)$ according to eq. (4.4) is a monotonously increasing function of $f_h(\theta)$ this result means that for m > 2 the aggregate total probability

$$P(X_h|\underline{0}) = \int_\theta P(X_h|\theta)p(\theta|\underline{0})d\theta$$

is always less than the probability arrived at by the methodology for crude binary choice, cf. section 4.1. Furthermore, this result explains the tendency towards over-estimation of the mobility proportions by the methodology for crude binary choice as mentioned in section 4.1.

4.3 Simultaneous Estimation

Unfortunately I have not succeeded in the effort of developing a methodology for ecological inference based directly on the model (4.2) for multiple choice. However, a methodology based indirectly on the model (4.2) has been developed by simultaneously considering multiple binary subsets of the voter's choice.

At election 1 we shall consider a genuine binary subset of the voter's choice, e.g. the possibility of voting for either party h or party m. It follows from the

model (4.2) for multiple choice that within this subset $(X_h UX_m)$, the probability for choosing party h is

$$P(X_h|(X_h UX_m),\theta) = \frac{P(X_h|\theta)}{P(X_h|\theta)+P(X_m|\theta)}$$

$$= \frac{\exp(\alpha_{oh}+\alpha_h\theta)}{\exp(\alpha_{oh}+\alpha_h\theta)+\exp(\alpha_{om}+\alpha_m\theta)}$$

$$= \frac{\exp[(\alpha_{oh}-\alpha_{om})+(\alpha_h-\alpha_m)\theta]}{1+\exp[(\alpha_{oh}-\alpha_{om})+(\alpha_h-\alpha_m)\theta]}$$

$$= \Psi(\alpha_o+\alpha\theta) \qquad (4.13)$$

where

$$\alpha_o = \alpha_{oh}-\alpha_{om}$$

and

$$\alpha = \alpha_h-\alpha_m$$

Analogously the probability at election 2 for choosing party j within the genuine binary subset $(Y_j UY_n)$ is

$$P(Y_j|(Y_j UY_n),\theta) = \Psi(\beta_o+\beta\theta) \qquad (4.14)$$

where

$$\beta_o = \beta_{oj}-\beta_{on}$$

and

$$\beta = \beta_j-\beta_n$$

Equations (4.13) and (4.14) show that within the tetrachoric subset

$$S: \ [(X_h \cup X_m) \cap (Y_j \cup Y_n)]$$

the model (4.2) for multiple choice is exactly equal to the model (3.3) for genuine binary choice.

This means that the methodology for binary choice can be applied to multiple choice situations, by simultaneously considering multiple tetrachoric subsets.

From table 4.1 we shall consider the genuine tetrachoric subset, that is, the combined possible outcome of voting for either X_h or X_m at election 1 and for either Y_j or Y_n at election 2. The proportions voting for each of the four combinations within a district are shown in the genuine tetrachoric table 4.2.

Table 4.2 Genuine Tetrachoric Table: Notation for Proportions of All Voters voting for either Party h or Party m at Election 1 and for either Party j or Party n at Election 2 within a District.

X: Choice at Election 1	Y: Choice at Election 2		
	j	n	Total
h	P_{hj}	P_{hn}	P_{hS}
m	P_{mj}	P_{mn}	P_{mS}
Total	P_{Sj}	P_{Sn}	P_S

Assuming that the consequences of all assumptions about the population distribution of θ apply to the subpopulation of all voters actually choosing within S^6, it follows that the tetrachoric correlation in table 4.2 can be estimated by the ecological correlation between the logits

$$x_{hm|jn} = \ln\left[\frac{P_{hS}}{P_{mS}}\right] \qquad (4.15)$$

$[P_{Sj}]$

across all districts.

The problem is, however: we do not know the marginal proportions P_{hS}, P_{mS}, P_{Sj}, P_{Sn} of table 4.2 for each electoral district. We only know the marginal proportions of table 4.1, that is the proportion of all voters voting for each party at the two elections in each election district.

The problem has found a solution by a computer iteration procedure. The first part of the procedure is to estimate the four proportions in table 4.2 within each district with the methodology for crude binary choice applying eq. (4.1):

$$\hat{P}_{hj} = G(P_{h.}, P_{.j}, r_{hj}) \quad ; \quad \hat{P}_{hn} = G(P_{h.}, P_{.n}, r_{hn})$$
$$\hat{P}_{mj} = G(P_{m.}, P_{.j}, r_{mj}) \quad ; \quad \hat{P}_{mn} = G(P_{m.}, P_{.n}, r_{mn}) \qquad (4.16)$$

With these 'first estimates' we compute the estimated marginal proportions

$$\hat{P}_{hS} = \hat{P}_{hg} + \hat{P}_{hn} \quad ; \quad \hat{P}_{Sj} = \hat{P}_{hj} + \hat{P}_{mj}$$
$$\hat{P}_{mS} = \hat{P}_{mj} + \hat{P}_{mn} \quad ; \quad \hat{P}_{Sn} = \hat{P}_{hn} + \hat{P}_{mn} \qquad (4.17)$$

within each district. These estimates are now taken as "data", and we compute the estimated "ecological" variables for genuine binary choice

$$\hat{x}_{hm|jn} = \ln\left[\frac{\hat{P}_{hS}}{\hat{P}_{mS}}\right]$$
$$\hat{y}_{jn|hm} = \ln\left[\frac{\hat{P}_{Sj}}{\hat{P}_{Sn}}\right] \qquad (4.18)$$

and the tetrachoric correlation is estimated by the "ecological" correlation across districts:

$$\hat{r}_{hj|mn} = r_e(\hat{x}_{hm|jn}, \hat{y}_{jn|hm}) \quad (4.19)$$

computed with eq. (3.37).

To estimate voter mobility between all parties in all districts it is not necessary to compute correlation coefficients for all possible permutations of binary choice. By choosing a reference-party X=m at election 1 and a reference-party Y=n at election 2 we only compute (4.19) for h=1,2,..., m-1 and j=1,2,..., n-1.

Because the consequences of the distribution assumptions for θ and μ are dependent on this choice, the final result of the estimation procedure depends upon which parties were chosen as reference-parties. The experience from experiments with different choices of reference-parties is that the most valid results are obtained, when the "neutral" party of non-voting is chosen as reference-party at each election.

Before the iteration procedure is initiated, first estimates of the proportions in the m'th row and the n'th column of table 4.1:

$$\hat{P}_{1n}, \hat{P}_{2n}, \ldots, \hat{P}_{mn} \quad ; \quad \hat{P}_{m1}, \hat{P}_{m2}, \ldots, \hat{P}_{m.n-1}$$

are computed with eq. (4.1).

An iteration procedure is now initiated which ensures that all (m-1)(n-1) correlation coefficients are simultaneously satisfied for all (m-1)(n-1) genuine tetrachoric tables involving the reference-parties, possible to pick from table 4.1.

The first step in the iteration procedure is to find "second estimates" of the mobility proportions

$$P_{hj} \quad ; \quad h=1,2,\ldots,m-1 \quad ; \quad j=1,2,\ldots,n-1$$

by the formula

$$\hat{\hat{P}}_{hj} = \frac{\hat{P}_{hn}\hat{P}_{mj}}{\hat{P}_{mn}} \frac{1+\hat{r}_{hj|mn}}{1-\hat{r}_{hj|mn}} \quad (4.20)$$

which is derived from equation (3.45).

These second estimates do not necessarily satisfy the logically expected relations:

$$\hat{\hat{P}}_{h.} = \hat{P}_{hn} + \sum_{j=1}^{n-1} \hat{\hat{P}}_{hj} = P_{h.} \quad ; \quad h=1,2,\ldots,m$$

$$\hat{\hat{P}}_{.j} = \hat{P}_{mj} + \sum_{h=1}^{m-1} \hat{\hat{P}}_{hj} = P_{.j} \quad ; \quad j=1,2,\ldots,n \quad (4.21)$$

So, the next step in the iteration procedure is to renew the first estimates involving the reference parties X_m and Y_n by

$$\hat{\hat{P}}_{hn} = \hat{P}_{hn} \frac{P_{h.}}{\hat{\hat{P}}_{h.}} \quad ; \quad h=1,2,\ldots,m-1$$

$$\hat{\hat{P}}_{mj} = \hat{P}_{mj} \frac{P_{.j}}{\hat{\hat{P}}_{.j}} \quad ; \quad j=1,2,\ldots,n-1 \quad (4.22)$$

and proceed with formula (4.20) to find third estimates of the mobility proportions, etc. until the relations (4.21) are satisfied with reasonable accuracy ($\pm 10^{-6}$ in this study).

Although the whole iteration procedure may seem very labourious it works very fast in the computer, even with more than 10 parties at each election.[7]

4.4 Estimation of Class Voting

The methodology for ecological inference was developed for estimation of voter mobility, but it was a gratifying experience to discover that the computer program for ecological estimation of voter mobility was able to estimate class voting as well.

If X_h instead of denoting "choosing party No. h at election 1" denotes "being a member of social class No. h at the time of election 2", it is possible to estimate the joint distribution of individual class-membership and party choice, using exactly the same computer program as for voter mobility. In this situation official census statistics about the class distribution at the time of election 2 is input to the computer program instead of the party distribution at election 1 for each district.

In my opinion this finding indicates that the latent structure approach which was developed for estimation of the covariation of voting behaviour between two elections, also applies to the covariation between class and party-choice. With this approach the explanation of the covarition between class and party is that the two variables have "something in common", i.e. the latent factors, as schematically displayed in figure 4.2.

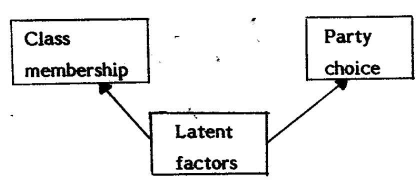


Figure 4.2 Schematic Presentation of the Latent Structure Approach for Estimation of the Covariation between Class and Party Choice

As mentioned in section 4.3 the simultaneous estimation of voter mobility between parties presupposes that a certain party is chosen as reference party, and that the most valid results was obtained, when the "neutral" party of non-voting was chosen at each election. It was a similar experience from experiments with different choices of a "reference-class" that the most valid results was obtained when the relatively "socially neutral" class of retired persons (pensioners) was chosen as reference-class.

Tables showing ecological estimates of voter mobility and class voting at Danish parliamentary elections 1920-1979 are dislayed in appendices 2 and 3. Some results from these tables are presented and discussed in the next chapter.

Notes to chapter 3

1. For very short time span between the occurrence of X and the occurrence of Y it will be necessary to deal with the problem of stochastic dependence not explained by the latent variables, but explained by the voters' memory of previous choice. The problem will not be discussed here because of the considerable time span usually existing between elections.

2. The assumption of constant θ on the individual level from one election to the other is a simplification. A more realistic "process" model assumes that the amount of change in θ is a function of the time span between the two elections. See also no. 1.

3. Note that the consequences of the assumption of isomorphism are independent of the arbitrarily assumed covariance matrix for the distribution (3.10) of μ (Thomsen, 1981).

4. A more detailed report of all mathematical derivations is given by Thomsen (1981).

5. More precisely, since the scale for α and β depends on the scale for κ, the convergence occur when both $\kappa\alpha\alpha' \to \infty$ and $\kappa\beta\beta' \to \infty$.

6. Admittedly this is a crude assumption, introduced with the purpose of applying the mathematical derivations obtained for genuine binary choice. Presumably the validity of this assumption depends on which parties are chosen of reference-parties, cf. the following text.

7. The FORTRAN computer program (Thomsen, 1981) can be requested from the author.

CHAPTER 4

ELECTORAL BEHAVIOUR IN DENMARK 1920-79

This chapter presents and discusses the validity of some ecological estimates of voter mobility and class voting at Danish parliamentary elections 1920-1979, selected from the tables in the appendices 2 and 3.

The estimates in these tables are based on data from "The Danish Statistical Commune Data Archive" (Bentzon, 1975, 1985; Madsen, 1986) at the Institute of Political Science. The geographical units in the archive are "communes" (local municipalities) with the exception of the city of Copenhagen, which is broken down to constituencies. To facilitate comparisons for the period 1920-1979, the units in the archive were grouped in 246 "constant units", i.e. geographical districts, almost identical throughout the period. Furthermore, the constant units were grouped into eight homogeneous political regions, proposed by Elklit (1978). With the exception of the City of Copenhagen each region consists of one or more counties, as indicated in table 0.1.

Table 0.1 Homogeneous Political Regions in Denmark

Region	Counties	Number of districts
1. City of Copenhagen	Copenhagen and Frederiksberg Municipalities	17
2. Copenhagen Suburbs	Københavns, Frederiksborg, Roskilde	41
3. Eastern Islands	Vestsjælland, Storstrøms, Bornholms	43
4. Funen	Fyns	25
5. East Jutland	Vejle, Århus	37
6. South Jutland	Sønderjyllands	23
7. West Jutland	Ribe, Ringkøbing	24
8. North Jutland	Viborg, Nordjyllands	36

The ecological estimation of voter mobility and class voting was carried out separately for each region by a FORTRAN computer program made by the author (Thomsen, 1981), and the results were cumulated for the whole country. It is these cumulated results which are displayed in the appendices 2 and 3.

The discussion of the validity of the ecological estimates takes the election of 1973 as a starting point. The election of December 1973 is outstanding in modern Danish political history, because "the four old parties", Social Democrats, Agrarian Liberals, Conservatives and Radical Liberals, which had dominated Danish politics since the turn of the century, all suffered setbacks and together lost 26 per cent of the votes. Instead two brand-new parties, the Progress Party and Centre Democrats, gained 24 per cent of the votes.

1 Ecological Estimation of Voter Mobility

When estimating voter mobility within a region the computer program for ecological inference takes as input from each district within the region the number of voters involved in two consecutive elections together with the percentage distribution of the parties at each election.

A problem in connection with the ecological estimation of voter mobility is the demographic transitions between the elections. Between elections young people get the franchise, other voters move from one district to another and others again die. A crude estimate of the number of dead voters is computed as a fraction of the number of old people in each district, and the increase of voters in each district plus those voters, who have replaced dead voters, was counted as "new voters". The outcomes of being a new voter or being dead are considered as "choice-alternatives" on par with party-voting and non-voting. Hence the "party"-distribution in each district includes as categories these awkward choice-alternatives.

The voter mobility is computed as the percentage of all voters moving from one party to another (which may be the same party). these results are shown in appendix 2 as table A for each election period. While the percentages in table A can be considered as "gross-movements" between pairs of parties, table B in appendix 2 shows the corresponding "net-movements" between pairs of parties. For instance the net-movement between Left Socialists and Communists is computed as the gross-movement from Left Socialists to Communists minus the gross-movement from Communist to Left Socialists. Finally, table C and D in appendix 2 show the gross-movements as respectively vertical and horizontal percentages, i.e. table C shows the distribution of party-choice at the first election given party-choice at the second election, while table D shows the distribution of party-choice at the second election given party-choice at the first election (dead voters excluded).

1.1 The validity of the estimates

Table 1.1 shows the estimated gross-movements from 1971 to 1973, i.e. the percentage of all voters involved in both elections moving between pairs of parties. Since one party, the German Minority Party in South Jutland, did not run in 1973, while two new parties, the Centre Democrats and the Progress Party did run in 1973, the number of choice-alternatives increased from 12 to 13. According to table 1.1, in spite of the substantial change in support for many parties, a large share of the voters voted for the same party at both elections. In fact, out of those voters that participated (i.e. actually voted) in both elections, a relative share of 62.4 per cent voted for the same political party.

Regarding the two new parties, the Centre Democrats received a great share of votes from former Social Democrats and from new voters. This also happened to the Progress Party, but besides receiving votes from former Social Democrats and new voters the Progress Party also received votes from the three other "old parties", Agrarian Liberals, Conservatives, and Radical Liberals. Furthermore the Progress Party received a considerable amount of votes from voters who abstained at the election in 1971.

Other significant examples of unstable voting behaviour are the gross-movements from Socialist People's Party to Communists, from Social Democrats to Socialist people's Party (and vice versa). According to the estimates a great number of former Social Democrats abstained at the election in 1973 and a significant share of non-voters from 1971 died before the election in 1973.

The high percentage of former German Minority voters voting for the Centre Democrats is due to the fact that the former German Minority candidate ran for the Centre Democrats. The very low per cent of stable Justice Party voters in spite of the increase in votes might be explained by the party's strongly novel position on the EEC-issue, the only non-socialist party taking a negative stance towards Danish membership.

To judge the validity of the ecological estimates they are compared with results from two interview panel studies that were made on representative national samples of Danish voters from 1971 to 1973. One was an academic study by Borre et al. (Borre, 1973, pp. 12-13; Worre, 1976, p. 21), the other was a study made by a private bureau Observa for a daily newspaper (Jyllandsposten, 1973).

l Ecological Estimates of Voter Mobility 1971-73. Per Cent of All Voters

	Election 1973													
	L.Soc.	Comm.	Soc.P. P.	Soc. Dem.	Cen. Dem.	Rad. Lib.	Just. P.	Chr. P.P.	Agr. Lib.	Cons.	Prog. P.	Non-Voting	Died 71-73	Total
	.5	.1	.1	.0	.1	.0	.1	.0	.0	.0	.1	.2	.1	1.3
	.0	.3	.3	.2	.1	.0	.0	.0	.0	.0	.0	.2	.0	1.1
	.1	1.3	2.7	1.5	.6	.0	.3	.0	.0	.2	.1	.6	.1	7.4
ı.	.2	.9	1.4	18.6	2.8	.1	.6	.2	.0	.6	1.2	2.9	.4	30.1
	.0	.0	.0	.1	.2	7.8	.3	.3	.9	.1	1.4	.2	.4	11.6
	.0	.0	.0	.0	.1	.1	.3	.1	.1	.0	.4	.2	.1	1.4
	.0	.0	.0	.0	.0	.0	.0	1.2	.0	.0	.1	.1	.0	1.6
	.0	.0	.0	.0	.1	.5	.0	.6	8.1	.0	2.7	.2	.5	12.6
	.1	.2	.3	.9	.8	.3	.4	.2	.4	6.4	2.4	.9	.2	13.5
.P.	.0	.0	.0	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.2
ing	.2	.2	.2	.5	.7	.3	.3	.5	.4	.2	2.5	5.4	1.1	12.5
:ers	.1	.0	.1	.2	1.1	.5	.2	.3	.5	.1	2.7	.6	.2	6.7
	1.3	3.1	5.1	22.0	6.6	9.6	2.5	3.5	10.5	7.8	13.6	11.4	3.1	100.0

Left Socialists
Communists
Socialist People's Party
ı. Social Democrats

Cen.Dem. Centre Democrats
Rad.Lib. Radical Liberals
Just.P. Justice Party
Chr.P.p. Christian People's Party

Agr.Lib. Agrarian Liberals
Cons. Conservatives
Prog.P. Progress Party
Germ.M.P. German Minority Party

Actually it is not an easy task to compare ecological estimates of voter mobility with interview surveys. For instance one problem is that voters are generally not willing to admit if they do not participate in an election, so usually abstainers are not included in published mobility tables from surveys. Another problem is that it is difficult to get in contact with as well very young as very old voters (cf. Damgaard, 1980, pp. 265-272), so the recorded voting behaviour among new voters and among those who had died during the election period tends to be unreliable. A third problem is that the party-distribution of the survey sample is usually not representative with regard to the actual election results unless a weighting procedure is applied.

My solution to the first two problems mentioned is to exclude abstainers, new voters and dead voters from the mobility tables when comparing ecological estimates with survey results. To the third problem my solution is to adjust the survey mobility tables (by multiplication of columns) so that the party-distribution at the second election is equal to the party distribution at this election among voters who did participate at the first election according to the ecological estimates. Of course, my solution to the third problem presupposes that the ecological estimates of the voting behaviour of former non-participants are valid. Of this I cannot be sure, but at least I think that the ecological estimates for former non-participants are more valid than the survey results, if they are at all published.

Table 1.2 shows the ecological estimates for the stable participants (taken from table 1.1) together with the adjusted mobility tables from the two survey panel studies. Notice that the bottom totals are the same in all three subtables, while this is not so for the right totals.

Compared with the survey results the validity of the ecological estimates seems to be quite good. Most of the conclusions reached at the beginning of this section are confirmed by the survey results. As mentioned, according to the ecological estimates the percentage of voters that voted for the same party among stable participants was 62.4, while this percentage is 65.4 in the Observa study and 60.4 in the Borre et al. study. As recorded by the ecological estimates, the Centre Democrats received a great share of votes from former Social Democrats and the Progress Party received votes from all four old parties, and these results are confirmed by the survey results. The survey results also confirms the mobility from Socialist People's Party to the Communists (The Borre study more so than the Observa study) and from Social Democrats to Socialist People's Party; but they do not confirm the mobility from Socialist People's Party to the Social Democrats.

where

$$TOT = \sum_{h=1}^{m} \sum_{j=1}^{n} P_{hj} = \sum_{h=1}^{m} \sum_{j=1}^{n} Q_{hj}$$

i.e. half of the cumulated absolute differences between equally positioned proportions in the two tables. The index, which has the minimum value 0 and the maximum value 1 can be interpreted as the amount of proportions which must be moved in one table to construct the other table. Applied to table 1.2 the amount of non-similarity between the ecological estimates and the Observa study is 16.2 per cent, while the amont of non-similarity between the ecological estimates and the Borre et al. study is 17.7 per cent.

It is interesting to note that the amount of non-similarity between the two surveys is 13.5 per cent. This result indicates that the similarity between the two surveys are only a little better than between the ecological estimates and the two surveys. Keeping in mind that the right totals in the tables with the survey results are not correct, it must be a fair judgement that the validity of the ecological estimates is at the same level as the interview results.

However, while the loss of validity concerning the survey results probably can be attributed to random sampling errors (i.e. to reliability) the loss of validity concerning the ecological estimates stems from systematic errors, probably because of the beforementioned "disturbance" from the rural-urban dimension.

1.2 Stability of Individual Voting Behaviour

The ecological estimates of voter mobility presented in appendix 2 offer the opportunity to explore the often advanced hypothesis that the stability of individual voting behaviour has declined during the century, and especially after the Second World War (cf. Converse, 1966; Worre, 1976).

Table 1.3 shows an index of individual stability. It is simply the ecologically estimated percentage voters choosing the same party at two consecutive elections out of all stable participants, i.e. voters who actually voted in both elections.

Table 1.3 Voter Stability. Voters Choosing the Same Party at Two Consecutive Elections. Per cent of Stable Participants[1]. Ecological Estimates

Election period	Stability	Election period	Stability
1920-1924	92.6	1953-1953	93.9
1924-1926	94.2	1953-1957	89.5
1926-1929	93.8	1957-1960	83.2
1929-1932	93.2	1960-1964	89.7
1932-1935	90.2	1964-1966	83.4
1935-1939	89.3	1966-1968	80.8
1939-1943	80.9	1968-1971	78.6
1943-1945	72.6	1971-1973	62.4
1945-1947	79.0	1973-1975	72.9
1947-1950	83.1	1975-1977	70.5
1950-1953	93.7	1977-1979	73.5

1) Stable participants: Voters participating (i.e. actually voted) in both elections. Computed from A-tables in appendix 2.

The stability is graphed in figure 1.1. As expected, a considerable decrease in stability is observed after the mid-fifties, but it is interesting to note that this came about after a sharp increase in stability since the election in 1945.

According to the ecological estimates, the stability was high in the twenties and thirties and then decreased in connection with the Second World War and the German occupation of Denmark. A stability as low as in the period 1943-45 was not reached again before the period in 1971-1973. Maximum stability (93.9 per cent) was attained in the short period from April to September between the two elections in 1953.

Of course the estimated amount of stability is chiefly explained by the amount of stability in the general support for the parties. But this is not the whole story, as we shall see in the next section.

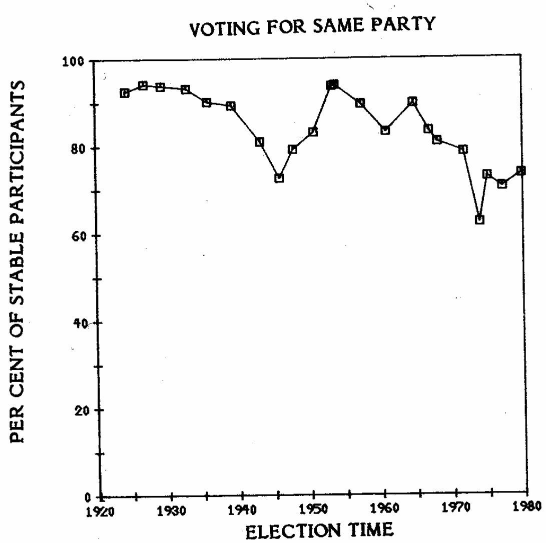


Figure 1.1 Voter Stability. Ecological Estimates (cf. table 1.3)

1.3 Net Movements between Pairs of Parties

As mentioned above, while the percentages in the A-tables in appendix 2 can be considered as "gross-movements" between pairs of parties, the B-tables show the "net-movements" between pairs of parties calculated by subtracting the mobility from a certain party Y to a certain party X from the mobility from X to Y.

This is done for two reasons. The first reason is that it is often more interesting to know the net "balance of payment" between two parties than the total amount of mobility between the two parties.

The other reason is that the ecologically estimated net-movements are probably less infected by the "disturbance" from the rural-urban dimension than the estimated gross-movements are. This assertion is reached by considering the tendency for overestimation of gross mobility within the group of rural parties and within the group of urban parties and the tendency for underestimation of gross mobility between the two groups, as mentioned in section 1.1. Since it is very likely that the bias, i.e. the tendency for either overestimation or underestimation of the mobility between two parties works in "both directions" (as well from party X to party Y as from party Y to party X), a considerable share of this bias is removed, when the net-movement from X to Y is considered.

It is important to note that if the same parties are competing at the two elections the B-tables can be read as well horizontally as vertically, since the tables are symmetric (with opposite signs) around the diagonal from upper left to lower right corner of each table. If for instance the table is read vertically, a column in the table represents the "net income" or gain of votes for a certain party in relation to all other parties.

A detailed examination of the tables for net-movements in appendix 2 (the B-tables) is outside the scope of this book, but some general observations are presented.

Concentrating on net movements greater than 1.0 per cent of all voters, the following pattern emerges.

If we at first look at the ecologically estimated net movements between pairs of political parties very little happens until 1943. Only once in the period from 1920 to 1939 did the net movement beween parties reach beyond the level of 1.0 per cent. That was in the election period 1932-1935 when the new agressive splinter party, Agrarians, gained at net support of 1.5 per cent from the Agrarian Liberals.

The period from 1939 to 1950 is a period with considerable net movements between political parties. Notable is the 1.2 per cent net gain in 1943 to the Social Democrats from the Communist Party which was made illegal by the German occupying power, and the very much amplified net movement in the reverse direction when the Communists shortly after the liberation from the occupation in 1945 gained 7.8 per cent from the Social Democrats. Once more the direction was reversed when the Social Democrats gained 4.1 per cent of the voters from the Communists in the election period 1945-1947. In this same period the controversial position of the Conservative leader probably explains the conservative loss of 1-2 per cent to as well the Social Democrats as to the Agrarian Liberals, which was reversed in the next election period 1947-1950, when the party had acquired new leadership.

After the relatively turbulent years from 1939 to 1950, the years from 1950 to 1966 were relatively quiet, concerning the estimated net movements between parties. Exeptions were the gain in 1960 of the new Socialist People's Party from the Communists (1.2 per cent) and from the Social Democrats (1.9 per cent), and again in 1966 the gain of the Socialist People's Party from the Social Democrats. This time 3.1 per cent, presumably because of dissatisfaction among former Social Democratic voters with their party's compromising position in relation to the non-socialist parties, for instance in the important housing question.

Since 1966 the net movements have been frequent, now also among the non-socialist parties. The sharply increasing popularity in 1968 of the Radical Liberals seems to have been paid by the Social Democrats (1.7 per cent) and the Agrarian Liberals (1.0 per cent). The "protest" election of 1973, which marked a widespread discontent with the achievements of "the four old parties", resulted in movements between parties that were unprecedented in Danish political history since 1920, as discussed in section 1.1. The remaining elections in the 1970's can to a certain extent be viewed as "reinstating elections" following the "turbulence" in 1973. According to the estimates, some of the 1973-gain of the Centre Democrats from the Social Democrats was paid back in 1975, but it is interesting to note that when the Centre Democrats surged in 1977, the party now drew votes from the non-socialist parties, most notably from the Agrarian Liberals. As a consequence of a short-lived government responsibility the Agrarian Liberals had a very successful election in 1975, mostly paid by the Progress Party, Radical Liberals and Conservatives, but the net gain from the Progress Party and from Conservatives was reversed in 1977. The tendency that the flow of votes in one direction at one election is reversed at the following election has also been observed for American presidential elections (Campbell, 1966).

Abstention

Since survey results are very unreliable concerning non-voting, the ecological estimates of voter mobility offer a particulary interesting opportunity for studying the role that abstention plays in the electoral process. The net movements from abstention to party-vote and reverse shall not be discussed in detail here, but the following general pattern emerges.

The first general finding is that party-victory is usually accompanied by net movement from abstention to the victorious party. Outstanding examples are the victories for the Social Democrats in 1924, 1929, 1943, 1960, 1977; for the Radical Liberals in 1968; for the Conservatives in 1943; and for the Agrarian Liberals in 1929, 1943, 1945, 1947, September 1953, and 1957. Keeping this general finding in mind, elections where a victory for a party is not followed by net movements from abstention to the victorious party becomes particulary interesting. Such examples are rare, but two call for special attention. The one is the victory for the Social Democrats in 1935, memorable because this was the best election result of all times for the Social Democrats. According to the ecological estimates it seems that the net gain for the party came neither from abstainers nor from other parties, but mostly from new voters. The other example concerns the victory for the Socialist People's Party in 1966, gained mostly with votes from the Social Democrats, but without substantial support from former abstainers.

The other general finding is the reverse to the first general finding, namely that election defeats are accompanied with net-movement from the defeated party to abstention. Clear examples are defeats for the Social Democrats in 1939, 1945, 1957, 1968, 1973; for the Conservatives in 1947; for Agrarian Liberals in 1935, 1950 and 1977; for Socialist People's Party in 1968; for Centre Democrats in 1975, and finally for the Progress Party in 1979. Exceptions to this general finding are a few cases where the loss from one party to a great extent were paid directly to another party, and thus did not contribute to abstention. Examples are the loss from the Communists to the Social Democrats in 1947, from the Social Democrats to the Socialist People's Party in 1966, and from the Radical Liberals to the Agrarian Liberals in 1975.

The two general findings are in accordance with the well established cross-pressure theory, that voters tend to abstain when confronted with conflicting forces (Lazarsfeld et al., 1944).

The Effect of Demographic Transitions

In appendix 2 persons outside the electorate are called "EJ VLG." (i.e. non-voters). At the first of the two elections in each table these are persons moving into the election district either by getting franchised or by geographical mobility, while at the second election they are voters from the first election who are now dead or have otherwise moved away from the election district. Thus the ecologically estimated net movement from a party or from abstention (i.e. non-voting) to non-voters or reverse can be regarded as the effect of demographic transitions, probably for the most part an effect of new generations getting the right to vote and at the same time older generations dying out.

Inspection of the B-tables in appendix 2 reveals a very uniform pattern concerning the effect of the continuous change in the generational composition of the Danish Electorate. At least until 1966 it is a uniform finding that the party which gained most advantage from the change in the electorate was the Social Democratic Party. Even in times of severe defeats as in 1939 and 1945, the Social Democrats received greater support from new voters than the party lost because old supporters died.

I think that this is the typical pattern for a social movement. Since 1966 the Social Democratic net gain from enfranchisement is still positive except at the election in 1973, but in most cases the D-tables reveal that the support for Social Democrats from new voters is less than from the rest of the electorate.

2 Ecological Estimation of Class Voting

When estimating class voting within a region exactly the same computer program as used for ecological estimation of voter mobility, takes as input for each district the number of voters at a certain election together with the percentage distribution of as well the social classes as the parties (including the "party" of abstention).

Since only one moment is involved there is no problem concerning the demographic transitions within the electorate as was the case when estimating voter mobility. However, now another problem is that the dates of the different censuses from which the social statistics are taken, does not coincide with the dates of the different elections. This problem is solved by estimating the class distribution in each district at the date of the election by linear interpolation between the census immediately before and the census immediately after the election in question (Thomsen, 1981).

The applied division in social classes reflects the important horizontal division of labour between agriculture and urban industries as well as the vertical division of labour between employers, white-collar employees and workers in the non-agricultural occupations. Furthermore, the two "non-active" classes of students and retired persons are included (Thomsen, op.cit.).

As for voter mobility the joint distribution of classes and parties is estimated in the computer as percentages of all voters, and the results are cumulated for the whole country. However, for substantial interpretations of the results the joint distribution as percentage of all voters is only of little interest. For this reason the results are displayed in appendix 3 as conditional percentage distributions. For each election table A shows "vertical percentages", i.e. the percentage distribution of the classes for each party, while table B shows "horizontal percentages", i.e. the percentage distribution of the parties (including abstention) for each class. Table C also shows horizontal percentages but excludes abstention from the distribution of the parties.

2.1 The Validity of the Estimates

Table 2.1 shows the estimated class voting at the election in December 1973 as per cent of each class (same as table 20.B in appendix 3).

The estimates indicate typical class-profiles for the different parties: The left-wing parties (Left Socialists, Communists, Socialist People's Party) are over-represented among white-collar employees and students, the Social Democrats among workers (non-agricultural). Among the non-socialist parties one group (Radical Liberals, Christian People's Party, Agrarian Liberals, Progress Party) is

Table 2.1 Ecological Estimates of Class Voting 1973. Per Cent of Each Class

s	L.Soc.	Comm.	Soc.P. P.	Soc. Dem.	Cen. Dem.	Rad. Lib.	Just. P.	Chr. P.P.	Agr. Lib.	Cons.	Prog. P.	Non-Voting	Total
iers	.0	.0	.0	.8	.6	15.6	.1	6.9	53.6	.3	20.0	2.0	100.0
ı Labourers	.1	.1	.2	2.7	.6	20.1	.2	9.0	45.2	.8	16.7	4.3	100.0
Employed	.5	.2	.4	4.4	3.9	15.0	1.2	6.9	16.3	9.2	34.7	7.3	100.0
e-Collar	2.3	3.5	7.6	24.2	9.5	8.5	4.2	1.5	3.3	13.8	11.8	9.7	100.0
:ers	.9	5.9	7.4	33.4	9.6	5.0	2.7	2.1	1.5	5.5	11.0	15.0	100.0
ents	12.5	6.0	10.8	13.2	6.4	7.4	6.2	1.8	4.9	13.9	6.1	10.8	100.0
ioners	[illegible]	1.3	2.8	22.1	2.9	14.2	1.1	5.8	16.6	5.6	11.6	15.5	100.0
l Electorate	1.3	3.2	5.3	22.7	6.8	9.9	2.5	3.6	10.8	8.1	14.0	11.7	100.0

c.	Left Socialists	Rad.Lib.	Radical Liberals	Prog.P.	Progress Party
m.	Communists	Just.P.	Justice Party		
?.P.	Socialist People's Party	Chr.P.P.	Christian People's Party		
Dem.	Social Democrats	Agr.Lib.	Agrarian Liberals		
Dem.	Centre Democrats	Cons.	Conservatives		

over-represented among agricultural occupations, while the rest (Centre Democrats, Justice Party, Conservatives) is over-represented among white-collar employees. The Progress Party is strongly over-represented among self-employed (in urban industries).

In the same way as already considered in section 1.1 regarding the mobility table 1971-1973, table 1.2, the validity of the ecologically estimated class voting is judged by comparison with results from two interview studies, made close to the election in 1973 on representative national samples of Danish voters. One was a study made by the bureau A.I.M. for a private publishing company (Boelsgaard, 1974) on a very large number of voters (N = 5,230). The period these data were collected (15.12.1973-15.5.1974) is over a considerable time span following the date of the election (4.12.1973). The other study was made by the Gallup Institute for an academic researcher (Glans, 1975; N = 1,766), 4-5 months after the election (April-May 1974).

It is easier to compare the results from these studies with the ecological estimates, than was the case concerning the mobility tables. However, since the amount of abstention is considered unreliable in the interview studies, abstention is excluded, when the ecological estimates are compared with the interview results. Furthermore, the party distribution in the interview studies are adjusted (by multiplication of columns in the table for the joint distribution of classes and parties), so that it is equal to the party distribution at the time of the election 1973.

Table 2.2 shows the ecological estimates of class voting for non-abstainers together with the adjusted tables from the interview studies. Notice that the bottom total distribution is the same for all three subtables (and equal to the official 1973-election results).

As was the case for voter mobility between parties, the ecological estimates of class voting are rather close to the survey results. Thus, most of the conclusions reached at the beginning of this section on the basis of the ecological estimates about the typical class profiles of the different parties are confirmed by the survey results. According to the survey results the left-wing parties are over-represented among white-collar employees and students (the Communists not in the A.I.M. study), the Social Democrats among workers. The same group of non-socialist parties (Radical Liberals, Christian People's Party, Agrarian Liberals, Progress Party) is over-represented among agricultural occupations, while the rest (Centre Democrats, Justice Party, Conservatives) is over-represented among white-collar employees. Finally, the Progress Party is strongly over-represented among self-employed.

Table 2.2 Class Voting according to ecological estimates and two interview studies. Abstention excluded

Class Voting 1973. Ecological Estimates. Per Cent of Class

Class	Y	K	F	A	M	B	E	Q	V	C	Z	Total
Farming	.0	.0	.1	1.2	.6	16.8	.1	7.4	53.5	.4	19.8	100.0
Self-Employed	.5	.3	.5	4.7	4.2	16.1	1.3	7.4	17.7	9.9	37.4	100.0
White-Collar	2.6	3.9	8.4	26.7	10.6	9.5	4.7	1.7	3.6	15.3	13.1	100.0
Workers	1.1	6.9	8.7	39.2	11.3	5.9	3.2	2.5	1.7	6.5	12.9	100.0
Students	14.4	6.7	12.0	14.7	7.1	8.3	6.9	2.0	5.5	15.6	6.8	100.0
Pensioners	.4	1.5	3.3	26.2	3.5	16.7	1.4	6.7	19.7	6.7	13.8	100.0
Total	1.5	3.6	6.0	25.6	7.8	11.2	2.9	4.0	12.3	9.2	15.9	100.0

Class Voting 1973. Survey Results (A.I.M.). Adjusted Per Cent of Class

Class	Y	K	F	A	M	B	E	Q	V	C	Z	Total
Farming	.0	.9	.0	1.7	5.5	18.1	1.0	6.0	47.7	1.5	17.7	100.0
Self-Employed	1.3	1.4	1.8	7.6	6.8	11.5	4.0	4.9	14.1	15.1	31.6	100.0
White-Collar	2.1	2.3	6.8	17.8	12.1	13.6	3.4	3.0	11.1	13.5	14.4	100.0
Workers	.8	6.6	8.4	38.6	7.9	6.7	2.8	2.4	4.5	2.9	18.4	100.0
Students	10.0	10.2	19.3	15.6	5.4	12.3	3.8	2.5	9.2	4.5	7.3	100.0
Pensioners	.0	1.8	2.2	38.8	2.8	11.5	2.0	7.0	15.3	11.0	7.6	100.0
Total	1.5	3.6	6.0	25.6	7.8	11.2	2.9	4.0	12.3	9.2	15.9	100.0

Class Voting 1973. Survey Results (Gallup). Adjusted Per Cent of Class

Class	Y	K	F	A	M	B	E	Q	V	C	Z	Total
Farming	.0	.0	.0	2.7	4.0	11.4	.0	5.6	45.3	7.5	23.4	100.0
Self-Employed	.0	.0	.0	10.6	5.9	14.4	7.5	1.2	11.8	16.5	32.1	100.0
White-Collar	3.4	3.6	7.9	17.4	9.8	14.5	2.7	3.8	8.2	13.6	15.1	100.0
Workers	1.2	5.1	8.4	36.3	10.4	8.4	2.9	2.7	5.0	3.2	16.1	100.0
Students	7.1	13.8	9.2	15.0	6.8	15.0	9.9	2.7	4.3	11.0	5.2	100.0
Pensioners	.0	2.9	4.7	38.9	3.9	9.4	1.6	6.9	13.4	10.8	7.5	100.0
Total	1.5	3.6	6.0	25.6	7.8	11.2	2.9	4.0	12.3	9.2	15.9	100.0

A: Social Democrats
B. Radical Liberals
C: Conservatives
E: Justice Party
F: Socialist Peopl. P.
K: Communists
M: Centre Democrats
Q: Christian Peopl. P.
V: Agrarian Liberals
Y: Left Socialists
Z: Progress Party

There is a marked difference between ecological estimates concerning the percentage of Social Democrats among retired persons (pensioners). This percentage seems to be underestimated by ecological inference, but this result might partly be due to the fact, that the high abstention among retired persons is not reliably recorded in the surveys.

A more serious argument against the validity of the ecological estimates is that a "disturbance" from the rural-urban dimension might appear similar to the one discussed in section 1.1. In fact, there seems to be a tendency towards that typical rural parties (B, Q, V) are overestimated among rural occupations (farming) while typical urban parties (Y, K, F, A, C) are overestimated among typical urban classes (white-collar, students) in relation to the survey results. This is, however, not a very strong tendency. Most notable is the strong underestimation of B, Q, and V among white-collar and students.

The index of non-similarity, presented in section 1.1, has been applied to compare the three subtables in table 2.2. Computed from tables showing the joint distribution of classes and parties as percentages of all voters (not presented here) the amount of non similarity between the ecological estimates and the A.I.M. survey is 16.7 per cent, and between the ecological estimates and the Gallup survey the amount of non-similarity is 16.1 per cent. Finally the amount of non-similarity between the two surveys is 12.2. These results are very much like the results obtained in section 1.1 when comparing ecological estimates of voter mobility with two different surveys. Taking in consideration that the data about class voting in the two surveys were collected up to five months after the election of December 1973, the conclusion is that the validity of the ecological estimates of class voting is at the same level as the survey results.

As for the ecological estimates of voter mobility the tendency that the loss of validity is systematic instead of random, can (as we shall see) be used to the advantage of the ecological estimates in connection with dynamic analysis.

In a recent research report Goul Andersen (1986b) has published the support for the socialist parties (the left-wing parties together with the Social Democrats) among four social classes as recorded in a number of interview studies since 1957. These results are shown in table 2.3 together with ecological estimates.

In table 2.3 it is interesting to note that in relation to the survey results the support for the socialist parties is overestimated among white-collar employees and underestimated among agricultural occupations (farming). Furthermore the ratio of over- and under-estimation seems to be rather stable across time. The estimates for the two other classes seem to be more realistic, especially the estimates of the working class support for the socialist parties.

Table 2.3 Class Voting for socialist parties according to ecological estimates and interview studies (Goul, 1986). Abstention excluded.

	Ecological Estimates				Interview Studies			
Election year	Farm-ing	Self-empl.	White-Collar	Work-ers	Farm-ing	Self-empl.	White Collar	Work-ers
1920	1.8	35.2	36.2	80.3	-	-	-	-
1924	2.6	40.2	38.6	83.2	-	-	-	-
1926	2.3	38.8	40.8	82.3	-	-	-	-
1929	3.5	45.4	51.6	86.0	-	-	-	-
1932	3.5	44.1	50.5	86.2	-	-	-	-
1935	4.1	45.6	51.4	86.6	-	-	-	-
1939	3.8	41.7	43.7	83.9	-	-	-	-
1943	3.6	42.1	43.6	79.6	-	-	-	-
1945	3.1	42.5	43.6	82.4	-	-	-	-
1947	3.0	39.1	48.7	86.5	-	-	-	-
1950	3.0	37.9	44.0	78.4	-	-	-	-
1953	2.3	35.6	46.1	79.9	-	-	-	-
1953	2.4	36.3	48.4	81.2	-	-	-	-
1957	1.5	30.2	48.5	75.8	-	-	-	73
1960	1.3	31.5	56.0	82.4	-	-	-	-
1964	1.5	29.5	51.8	79.5	5	26	39	79
1966	2.0	28.4	51.5	80.1	5	22	41	82
1968	1.8	20.9	44.0	72.0	4	15	36	72
1971	1.8	17.8	53.5	76.6	4	20	40	76
1973	1.3	6.0	41.6	55.9	2	12	30	54
1975	1.6	6.7	47.6	65.1	2	12	37	63
1977	1.5	12.1	56.2	69.2	1	14	45	64
1979	1.8	14.3	58.0	72.7	7	21	50	66

The fact that the bias (the tendency towards either over-estimation or under-estimation) is stable across time makes it particularly interesting to study the change across time of the class support for the different parties. Since the bias is stable, the change of support is presumably more valid estimated than the support in a certain moment. Only a few examples shall be presented, viz. the support for "the four old parties" 1920-1979.

2.2 Class Support for "the Four Old Parties"

Ecological estimates of historical class voting for "the four old parties", i.e. the four parties having seats in the Danish Parliament (Folketing) throughout the whole period 1920-1979 are graphed in figures 2.2, 2.3, 2.4, 2.5. The figures show stable class-profiles for the parties. With few exceptions each party has a certain rank order of support in the different classes throughout long periods.

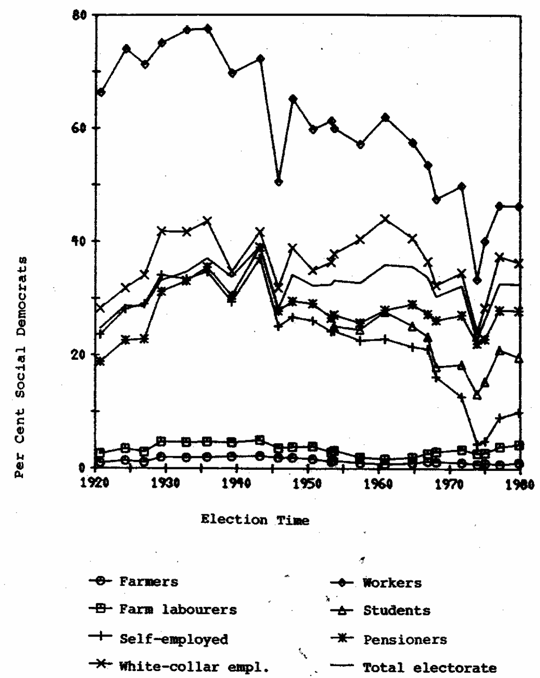


Figure 2.2 Ecological Estimated Class voting for Social Democrats

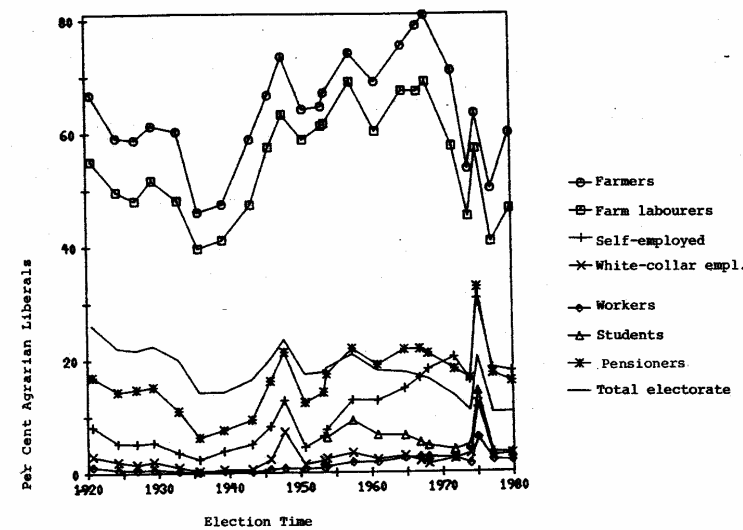


Figure 2.3 Ecological Estimated Class Voting for Agrarian Liberals

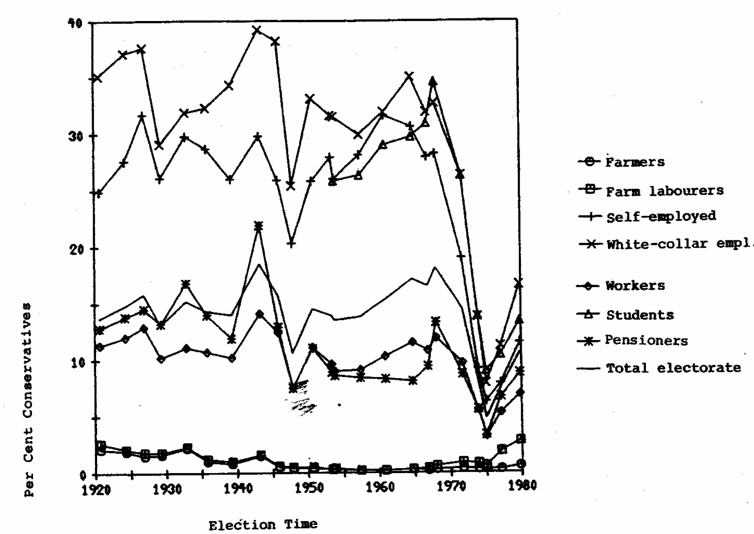


Figure 2.4 Ecological Estimated Class Voting for Conservatives

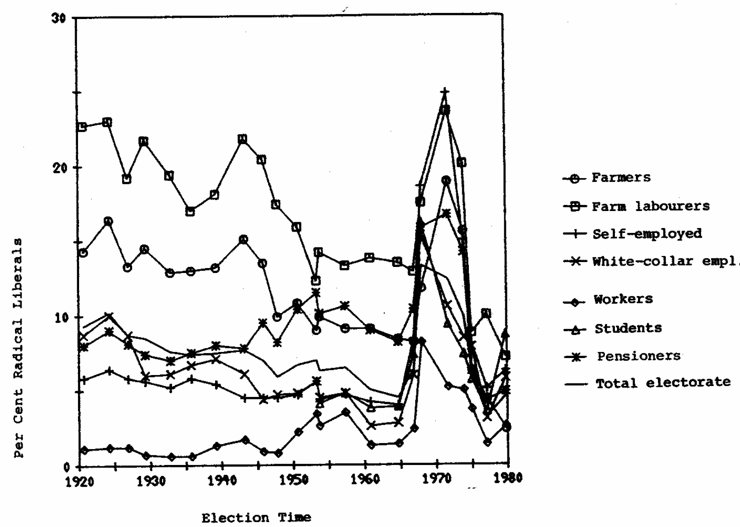


Figure 2.5 Ecological Estimated Class Voting for Radical Liberals

The Social Democratic Party has more than any other party participated in government formation. The government responsibility which necessitated co-operation with centrist parties, especially the Radical Liberal Party, seems to have weakened the strong working-class support from the between war years. Even though the party since the Second World War had lost support in most classes (especially among self-employed) it has kept a relatively high proportion of the total electorate. This paradox is explained by the fact that the change in the Danish class distribution (especially the growing "new middle class" of white-collar employees) has been to the advantage of the Social Democratic Party.

The opposite phenomenon holds for the Agrarian Liberals. Although the party seems to have maintained support among voters from agricultural occupations and even gained support among self-employed, it has lost support in the total electorate at most elections. Exceptions are the two "charismatic" elections in 1947 and 1975 where strong party leaders gained support from classes not traditionally voting for Agrarian Liberals.

Until 1968 the Conservative Party was the typical party for the urban middle-class, i.e. white-collar employees and self-employed. In the following elections the

party's strength was badly weakened by conflicts in the leadership of the party. However, the party recovered after 1975.

Traditionally, the Radical Liberal Party co-operated with the Social Democrats and secured support from farm labourers and smallholders but also from some segments of the urban middle-class (e.g. school-teachers and intellectuals). A marked change in party-image, accomplished by a new party leader in the mid-sixties with the proposal of more co-operation with the non-socialist parties, produced a series of unprecedented elections results. However, the leadership in the 'bourgeois government' 1968-1971 (majority government by the Radical Liberals, Agrarian Liberals and the Conservatives) resulted in very different reactions from the different classes. From 1968 to 1971 the Radical Liberals gained considerable support from especially self-employed and agricultural occupations, but lost support from white-collar employees, workers and students. Since 1971 the party seems to be weakened by discussions about "party-identity".

3 CLASS POLARISATION

In Danish Politics, 1920 and 1979 can be regarded as opposite extremes in the historical development from strong class-antagonism to peaceful class-relations in the Danish society. The 1920 election found the labour class movement in strong opposition to the bourgeois establishment which had supported the anti-parliamentarian maneuvres from the king in the same year, while in 1979 it was the first time that the political "branch" of the labour class movement, the Social Democratic Party, had experienced a governmental alliance with the main spokesman for the non-socialist parties, the Agrarian Liberals.

In table 2.6 and 2.7 the ecologically estimated class support among workers and non-workers for the socialist parties is shown for the two elections 1920 and 1979 respectively (the percentages are cumulated from table 1.C and 23.C in appendix 3).

Table 2.6 Class support for socialist parties, 1920. Pct. of valid votes

	Social. parties	Non-soc. parties	
Workers	80.3	19.7	100.0
Non-workers	18.8	81.2	100.0

Table 2.7 Class support for socialist parties, 1979. Pct. of valid votes

	Social. parties	Non-soc. parties	
Workers	72.7	27.3	100.0
Non-workers	42.1	57.9	100.0

The weakening of class antagonism in Danish society is clearly reflected when comparing the two tables. The tables indicate that the support for the socialist parties has decreased within the working class, while it has increased within the the rest of the society.

There are several ways of measuring the amount of class-antagonism or "class-polarisation" in tables such as table 2.6 and 2.7, and an extensive litterature discusses this problem (cf. Alford, 1963; Korpi, 1973; Lipset, 1981; Glans, 1983; Goul Andersen, 1986a). Most usually practiced is the socalled Alford-Index, which computes the difference between the socialist support in the two groups. Utilizing this index the class polarisation is

80.3 - 18.8 = 61.5

in 1920, and

72.7 - 42.1 = 30.6

in 1979; clearly a considerable decrease in class polarisation as also should be expected from the historical experience. The problem is, however, that other indexes give other results concerning the amount of decrease, and it is difficult to judge which index is the most appropriate.

In appendix 1, section 3, an index for class polarisation is derived from the ecological linear logit model, presented in chapter 2. According to this model, the proportion P_{2t} in class No. 2 (non-workers) supporting the socialist parties should not be subtracted from the proportion P_{1t} in class No. 1 (workers) supporting the socialist parties. Instead the proportions should be logit-transformed before subtracted, i.e. the index for class polarisation, called "the logit difference" is

$$\text{Logit Difference} = \ln\left[\frac{P_{1t}}{1-P_{1t}}\right] - \ln\left[\frac{P_{2t}}{1-P_{2t}}\right]$$

One reason for preferring the logit difference to the proportion difference is that the logit difference is not sensible to "floor" and "ceiling" effects on proportions close to either 0 or 1 as discussed in chapter 2, section 1. Another reason is that the model from which the logit difference is derived is equal to the Rasch model of measurement, when the logit difference is constant across time. As

implied from the discussion in chapter 2, section 2.2, unchanged logit difference means that class polarisation in a genuine (mathematical statistical) sense is independent of elections.

It must be mentioned that the logit difference is the logaritmic transformation of the odds ratio:

$$\text{Odds Ratio} = \frac{P_{1t}/(1-P_{1t})}{P_{2t}/(1-P_{2t})}$$

which is extensively applied as an index for covariation in four-fold tables in many research fields (cf. Fleiss, 1973; Plackett, 1974; Breslow and Day, 1980).

Table 3.3 Class Index. Unstandardized and standardized. Ecological Estimates

Election Year	Socialist Support: Workers	Socialist Support: Middle Class Unstandardized	Socialist Support: Middle Class Standardized	Class Index Unstandardized	Class Index Standardized
1920	80.3	18.8	22.6	2.9	2.6
1924	83.2	22.6	26.2	2.8	2.6
1926	82.3	23.5	27.2	2.7	2.5
1929	86.0	28.7	32.9	2.7	2.5
1932	86.2	29.5	33.2	2.7	2.5
1935	86.6	32.7	35.9	2.6	2.4
1939	83.9	29.1	31.3	2.5	2.4
1943	79.6	30.6	31.9	2.2	2.1
1945	82.4	29.2	30.1	2.4	2.4
1947	86.5	29.6	30.3	2.7	2.7
1950	78.4	29.1	29.1	2.2	2.2
1953	79.9	29.2	28.8	2.3	2.3
1953	81.2	29.7	28.8	2.3	2.4
1957	75.8	28.4	26.4	2.1	2.2
1960	82.4	34.2	30.6	2.2	2.4
1964	79.5	34.4	28.5	2.0	2.3
1966	80.1	35.8	28.6	2.0	2.3
1968	72.0	30.5	23.8	1.8	2.1
1971	76.6	37.8	26.2	1.7	2.2
1973	55.9	29.1	18.4	1.1	1.7
1975	65.1	32.0	19.8	1.4	2.0
1977	69.2	39.3	24.3	1.2	1.9
1979	72.7	42.1	25.8	1.3	2.0

Table 3.3 shows the support for the socialist parties in the working class and among non-workers respectively. The support among non-workers is computed as well unstandardized as standardized. The unstandardized support among non-workers for the socialist parties is simply computed as the percentage support of all valid votes among non-workers regardless of the size of the different social classes (i.e. farmers, self-employed, white-collar employers etc.) among non-workers. In contradistinction to that the standardized support is measured by weighting the support from each subclass with the same weight throughout the whole period 1920-79. As weights for the different classes was chosen the relative proportion of each class in 1950. With this procedure the standardizd socialist support is computed as if the internal class-composition among non-workers was unchanged (and the same as in 1950) for the whole period 1920-79.

Now the class polarization in Denmark is measured as well unstandardized as standardized by the logit difference between the support in the working class and the support among non-workers, either unstandardized or standardized.

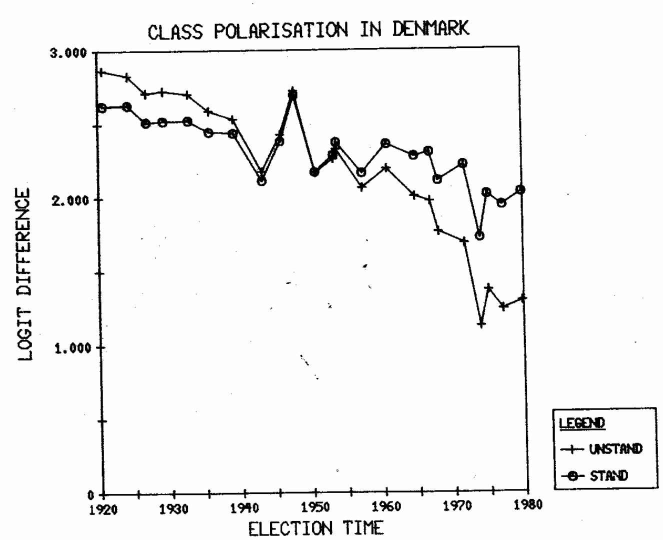


Figure 3.1 Unstandardized and standardized Class Index (cf. Table 3.3)

In figure 3.1 the index for unstandardized class polarisation is graphed together with the index for standardized class polarisation. The curves in figure 3.1 show that the decrease of class polarisation in Denmark is much steaper when the unstandardized index is considered instead of the standardized.

The interpretation of these results could be that the decrease of class polarisation is as well attributed to decreasing psychological antagonism between the working class and the middle class as to the changing class structure of the Danish society (cf. Goul Andersen, 1986a).

It is interesting to note that the standardized curve in the period 1920-79 is only (but not always) decreasing if the government is headed by the Social Democratic Party as was the case in 1924-26, 1932-35, 1935-39, 1939-43, 1947-50, 1953-57, 1960-64, 1966-68, 1971-73, and 1975-77, while the curve is never decreasing if the government is non Social Democratic as in 1920-24, 1926-29, 1945-47, 1950-53, 1968-71, and 1973-75.

This finding indicates that Social Democratic governments have played an important role in the decrease of class-antagonism in this century (cf. Thomsen, 1986).

CHAPTER 5

SUMMARY

The Problem of Ecological Inference

The main subject of this book is an attempt to solve the problem of ecological inference, i.e. the practice of inferring individual-level behaviour from aggregate level behaviour, within the research field of electoral studies.

In practice the aim of electoral ecological inference is to use official statistics from geographical areas to answer questions such as: "How many voters changed their support from party X to party Y from election 1 to election 2?", and "How many voters in class A voted for party X?", without conducting expensive surveys. Especially in historical electoral research the need for valid methods for ecological inference is great, simply because we often do not have information on past individual behaviour.

During the 1950s and 1960s the practice of ecological inference was almost banned as an indecent methodological fallacy. Since then, much work has been done to stipulate under which conditions ecological inference is actually possible (cf. chapter 3, section 1).

Unfortunately, as discussed in chapter 3, section 2, to this day electoral ecological analysts have not been very successful in their attempts to estimate individual voting behaviour from ecological data. As argued by the present author, one reason for this lack of success is the predominant use of regression methods, which seems to be ridden by the two evil bedfellows called "the problem of mis-specification" and "the problem of multicollinearity" (cf. chapter 3, section 2.3), as soon as the number of political parties or social classes involved is greater than two.

Instead of the seemingly futile regression approach I propose another approach based on latent structure theory. Latent structure theory was developed within the fields of attitude measurement and psychological testing, but it is my firm belief that the theory has important possibilities of application also for many other human behaviour research fields.

What led me on the track of latent structure theory as a vehicle for ecological inference was the theoretical finding, presented in chapter 3, section 3, that the logit model I had applied for analysis at the aggregate (ecological) level could, at least as a fair approximation, be mathematically derived from a latent structure model at the individual level.

The Ecological Model for Voting Behaviour

As presented in chapter 2, section 1, the starting point for the development of a general ecological model for voting behaviour is the empirical finding, that the pattern of change in the support for political parties across election districts tends to be "multiplicative" for small parties and "additive" for big parties. Instead of using different linear models depending of the size of the party investigated I propose a general non-linear model (the constant odds ratio model) for uniform swing which is approximately multiplicative for small parties and approximately additive for big parties. The advantage to this model is that it defines uniform swing independently of the amount of support for the party.

In fact, the reason for the choice of the constant odds ratio model is not only its suitability as an instrument for data description. In this respect other models such as the linear proportion model work quite as well. As discussed in chapter 2, section 2.2, from the point of view of mathematical statistics, the real strenght of the constant odds-ratio model is that it is the one and only model (for binary data) that, if valid, describes the covariation between election districts and parties independent of elections, and thus is a very appropriate model for constant social orientation of a party across time. The explanation why uniform swing is predominant in many political systems is that political parties normally have a fairly stable social orientation (e.g. position on class-issues) while their general popularity may change considerable from one election to the next.

With the purpose of explaining deviations from uniform swing in chapter 2, section 2, a multi-dimensional linear logit model is proposed which allows for change in the social orientation of a party. This model has the model for uniform swing as a special case when the social orientation of the party is constant across time. Using the linear logit model an important finding from the Danish election project is, that the development in the social orientation of the Danish political parties 1920-1979 is estimated as being much more stable than if an ordinary linear proportion model is used. This is the main reason for prefering the more complex logit model instead of an ordinary linear proportion model, although the ordinary linear proportion model describes the covariation of the ecological data just as good as the logit model.

A formal weakness concerning the linear logit model is that it applies exclusively to binary choice, i.e. it does not take the complex competition among parties in multiparty systems into account. For this reason, a linear multiple logit model is proposed in chapter 2, section 3, which has the binary choice model as a special case. However, although the multiple choice model is more correct from a formal point of view, in the Danish election project it does not give results that are very different from the results obtained with the binary choice model. This finding indicates that for this kind of ecological data the binary model is a reasonable approximation to the multiple choice model.

The Latent Structure Theory for Binary Choice

As implied in the first section of this summary the main problem concerning ecological inference is to specify under what conditions individual behaviour can be inferred from ecological (aggregate) behaviour. The mathematical side of this problem is to formulate mathematical models for individual behaviour and then derive the consequences of these models for the aggregate level, given certain assumptions about the relation between the individual and the ecological variation. Furthermore, it is important for the possibility of ecological inference that it is also mathematically possible "to go in the opposite direction", i.e. from the aggregate level to the individual level.

Of course it is not enough to construct mathematical theories which in principle allow for cross-level inference. It is equally important to consider if the models for behaviour inherent to a mathematical theory applies to actual behaviour.

Since the linear logit model for binary choice served so well for ecological analysis of actual electoral behaviour it was natural to look for at mathematical theory from which the ecological linear logit model could be derived. Unfortunately, as mentioned in chapter 3, section 3.1, it turned out that if the multidimensional linear logit model was proposed at the individual level, this could not lead to the derivation of an analogous multidimensional linear logit model at the ecological level. However, if the multidimensional linear probit model, which is a very close approximation to the logit model, is used instead at the individual level, it is actually possible to derive an analogous probit model at the ecological level (cf. chapter 3, section 3.3).

To make ecological inference possible with this probit theory, it is necessary to introduce certain assumptions about the relation between the latent structure at the individual level and the latent structure at the ecological level. Under the assumption of isomorphism, i.e. the covariance-matrix of the individual level is proportional to the covariance-matrix of the ecological level; and the additional assumption of a very high ratio of the variance at the individual level in relation to the variance at the ecological level it turns out that the total individual tetrachoric correlation is approximately equal to the ecological probit correlation (cf. chapter 3, section 3.5).

The two assumptions, i.e. the assumption of isomorphism and the assumption of high variance ratio, are only believed to be approximately valid within homogenous political regions where the ecological variation between election districts is of the same "nature" (i.e. has the same dimensional structure) as between individual voters (cf. op.cit.). Furthermore, since the ecological logit correlation is a close approximation to the ecological probit correlation, the hypothesis behind the ecological estimation of individual voting behaviour is as follows:

> **For binary choice within a homogenous political region the total individual tetrachoric correlation is equal to the ecological logit correlation.**

This result is a bit confusing, since as well the advocates as the critics of the practice of ecological inference unanimously agree in stating that the individual correlation is seldom equal to, but usually weaker than the ecological correlation (cf. chapter 3, section 2.3). The confusion can to some extent be removed by considering the fact that it is usually the Pearson product-moment correlation, which is computed at the individual level, even in cases where the correlated variables are binary. Thus it turns out that the most quoted example of "the ecological fallacy" originally put forward by the harsh critic of ecological inference, Robinson, concerning the correlation between colour and illiteracy actually shows that the individual correlation is approximately equal to the ecological correlation, if the tetrachoric correlation instead of the product-moment correlation is computed at the individual level.

Ecological Inference for Multiple Choice

As mentioned in connection with the discussion of ecological models it was found in the Danish election project that the binary choice model is as good an instrument for ecological analysis of multiparty voting behaviour as a more complex multiple choice model. However, this is not so when the level of description is the individual voters. When using the hypothesis from the last section for direct ecological estimation of individual voting behaviour in Denmark it turned out that there was a clear tendency towards overestimation of voter mobility between parties (chapter 3, section 4.1). The explanation of this tendency was that the party choice was considered as a "crude" binary choice between a certain party and all other parties instead of a "genuine" binary choice between only two different parties (section 4.2).

This was a serious obstacle for electoral application of the theory for binary choice because even in two-party systems the option of abstention is a possible third choice-alternative. At this stalemate in the research process the previously renounced multiple logit model showed its strength. Although it was not possible to use this model for direct ecological estimation of individual voting behaviour, the model showed that an indirect approach was possible by simultaneous considering different binary subsets of the voter's choice.

The problem of simultaneous considering many different binary subsets was solved by a computer iteration procedure. With this procedure the hypothesis above for genuine binary choice is applied by simultaneously considering all possible binary choices, each involving a specific "reference-party" (e.g. the "neutral" party of non-voting) and one of the other parties. By the iteration procedure the computer then gradually constructs all possible tetrachoric (four-fold) tables which simultaneously satisfy the hypothesis of identical individual and ecological correlation, and the voter mobility between all parties are estimated.

Although the estimation procedure is developed on basis of a theory for voter mobility between two elections, it turns out that exactly the same computer program can estimate class voting by substituting ecological data for the different parties at one election with ecological data for the different social classes.

Ecological Estimation of Voting Behaviour in Denmark

Ecological estimates of voter mobility and class voting at Danish parliamentary elections 1920-1979 are presented in appendices 2 and 3. In chapter 4 the validity of the ecological estimates is evaluated by comparing the estimates of voter mobility between the elections in 1971 and 1973 with two interview panel studies and by comparing the ecological estimates of class voting in 1973 with two different surveys. The conclusion from these investigations is that the validity of the ecological estimates is on the same level as the validity of the interview results.

However, it is important to note that the loss of validity of the ecological estimates tends to be caused by systematic errors rather than by random errors, which is contrary to what is customary with regard to survey results. This tendency towards systematic errors can be used to the advantage of the ecological estimates in connection with dynamic analysis. Since the systematic bias (the tendency towards either over-estimation or under-estimation of percentages) seems to be rather stable across time, the bias is to a certain extent "cancelled out" when the amount of change is studied.

One example is that the "gross-movements" between parties tends to be overestimated within the group of rural parties and within the group of urban parties, while the gross-movements between these two groups tend to be underestimated. However, when "net-movements" between pairs of parties are considered, some of the bias is presumably removed.

The same kind of reasoning applies when studying the change in "the class profile" of a certain party or group of parties. Although the ecologically estimated support for the party or the group of parties might be biased at a certain moment, the change in support is probably less biased.

Applying this "dynamic principle" to the tables in the appendices 2 and 3 the results seem to be quite valid. Especially the net movements between pairs of parties seems to be in accordance with what is to be expected from the recorded Danish political history.

In chapter 3, section 3, an index of class polarisation, derived from the model for ecological analysis, is applied to the ecologically estimated socialist vote within the working class and the middle class respectively. It might be an indirect proof of the dynamic validity of the ecological estimates that by applying this index a very clear connection is found between the political composition of the government in the election period and the change in class polarisation during the same period.

Further Development of the Methodology for Ecological Inference

In my opinion the relatively high validity of the results obtained with the presented methodology for ecological inference suggests the superiority of the latent structure approach in relation to the contemporary predominant regression approach. This is especially so considering the invalid results obtained with the regression approach, if the ecological estimation of individual voting behaviour concerns more than two parties or more than two social classes (chapter 3, section 2). In contradistinction to that, the number of parties or classes presents no problem for the methodology based on the latent structure approach.

The methodology presented in this book is of course not the final answer about how to apply the latent structure approach to ecological inference in electoral research. In fact, the two selected assumptions within the latent structure logit-theory, viz. the assumption of isomorphism and the assumption of very high variance ratio between the individual and the ecological level, represents rather extreme simplifications. For instance, the assumption of isomorphism seems not to be quite valid, considering that the rural-urban dimension tends to influence the ecological estimates too much in relation to the middleclass - workingclass dimension.

This finding suggests that more valid estimates can be obtained with the latent structure theory by formulating more sophisticated assumptions about the relation between the individual and ecological level - a matter for further development of the methodology for ecological inference.

CHAPTER 6

DANSK SAMMENDRAG

En logit tilgang til økologisk analyse og inferens vedrørende politisk vælgeradfærd

Denne afhandling skal ses som et forsøg på at integrere de metodologiske resultater, jeg er nået frem til i mit arbejde med det danske valgprojekt ved Institut for Statskundskab siden 1970erne. Projektet vedrører økologisk analyse og inferens i forbindelse med officiel statistik fra folketingsvalg og folketællinger siden 1920. Eftersom den metodologi for økologisk analyse, jeg har udviklet, er udførligt beskrevet andetsteds (Thomsen, 1971, 1972, 1979) er afhandlingens hovedemne et forsøg på inden for vælgeradfærdsforskningen at løse problemet om økologisk inferens, dvs. spørgsmålet om, hvorvidt man kan slutte sig til adfærd på individ-niveau ud fra adfærd på aggregeret niveau.

Problemstillingen vedrørende økologisk inferens

I praksis er formålet med økologisk inferens at anvende offentlig statistik opgjort på geografiske områder til at besvare spørgsmål så som: "Hvor mange vælgere gik fra at stemme på parti X til at stemme på parti Y fra valg nr. 1 til valg nr. 2?" eller: "Hvor mange vælgere i social-klasse A stemte på parti X?", uden at skulle gennemføre bekostelige interview-undersøgelser. Særligt inden for historisk valgforskning er der stort behov for gyldige metoder til gennemførelse af økologisk inferens, simpelthen fordi man ofte ikke har oplysninger om tidligere individuel adfærd.

Gennem 1950erne og 1960erne var økologisk inferens nærmest bandlyst som en uanstændig metodologisk fejlslutning; men siden da har der været udført et stort arbejde for nærmere at fastlægge under hvilke betingelser økologisk inferens faktisk er mulig (jvf. kapitel 3, afsnit 1).

Indtil nu har forskere i politisk økologi (som diskuteret i kapitel 3, afsnit 2) desværre ikke haft særligt held i deres forsøg på at estimere individuel vælgeradfærd ud fra økologiske data. Efter min mening er en af grundene til den manglende succes den udbredte anvendelse af regressions metoder, der ser ud til at være forfulgt af de to ondskabsfulde plageånder, der hedder "misspecifikation" og "multicollinearitet", lige så snart antallet af involverede politiske partier eller sociale klasser er større end to.

I stedet for den tilsyneladende ufrugtbare regressions tilgang foreslår jeg en anden fremgangsmåde, der er baseret på latent struktur teori. Latent struktur teorien er udviklet med henblik på holdningsmåling og psykologisk testning; men jeg er overbevist om, at teorien også har vigtige anvendelsesområder inden for mange andre samfundsvidenskabelige forskningsområder.

Det, som ledte mig på sporet af latent struktur teorien som et redskab for økologisk inferens, var den teoretiske opdagelse, vist i kapitel 3, afsnit 3, at den logit model, jeg havde anvendt til analyse på det aggregerede (økologiske) niveau, kunne, i det mindste som en rimelig approksimation, matematisk udledes fra en latent struktur teori på det individuelle niveau.

Den økologiske model for vælgeradfærd

Som vist i kapitel 2, afsnit 1, var udgangspunktet for udviklingen af en generel økologisk model for vælgeradfærd den empiriske konstatering, at forandringsmønstret for tilslutningen til politiske partier i en række valgdistrikter er tilbøjelig til at være "multiplikativ" for små partier og "additiv" for store partier. I stedet for at være henvist til at bruge forskellige lineære modeller, alt efter størrelsen af det undersøgte parti, foreslår jeg en generel ikke-lineær model (den konstante odds rate model) for ensartet udsving, som er tilnærmet multiplikativ for små partier og tilnærmet additiv for store partier. Fordelen ved denne model er, at den definerer ensartet udsving for et parti uafhængigt af den samlede tilslutning til partiet.

I virkeligheden er begrundelsen for valget af den konstante odds rate model ikke kun dens velegnethed som et instrument for databeskrivelse. Hvad dette angår, er andre modeller, såsom den lineære andels model, ofte lige så anvendelige. Som diskuteret i kapitel 2, afsnit 2.2, er den virkelige styrke ved den konstante odds rate model, set fra den matematiske statistiks synsvinkel, at det er den eneste model (for binære data), som, hvis den er gyldig, kan beskrive samvariationen mellem valgdistrikter og partier uafhængigt af de enkelte valg og således er en meget velegnet model for den situation, hvor et politisk partis sociale orientering er uændret over tid. Forklaringen på, at ensartede udsving er så almindelig i mange politiske systemer, er, at politiske partier gennemgående har temmelig stabil social orientering (for eksempel hvad angår holdning i klasse-spørgmål), mens det enkelte partis samlede popularitet kan ændre sig betragteligt fra det ene valg til det andet.

Med det formål at kunne forklare afvigelser fra ensartede udsving foreslås i kapitel 2, afsnit 2, en multidimensional lineær logit model, som tillader, at et partis sociale orientering kan ændre sig over tid. Denne model har modellen for ensartet udsving som specialtilfælde i den situation, hvor et partis sociale orientering er uændret. Et vigtigt resultat i det danske valgprojekt er, at udviklingen i danske partiers sociale orientering 1920-79 skønnes at være langt mere stabil, hvis den lineære logit model anvendes i stedet for den lineære andels model. Dette er den vigtigste grund til at foretrække den mere komplekse logit model frem for den simple lineære andels model.

En formel svaghed ved den lineære logit model er, at den strengt taget kun kan anvendes ved binært valg, da den ikke tager hensyn til den komplekse konkurrence mellem partier i et flerparti-system. På denne baggrund foreslås i kapitel 2, afsnit 3, en lineær multipel logit model, der har modellen for binært valg som specialtilfælde. Det viser sig imidlertid i det danske valgprojekt, at skønt modellen for multipelt valg set fra en formel synsvinkel er mere korrekt, giver den ikke resultater, der er meget forskellige fra de resultater, der er opnået med modellen for binært valg. Denne erfaring viser, at for denne type økologiske data er den binære model en rimelig approksimation til den multiple model.

Latent struktur teorien for binært valg

Som det fremgik af det andet afsnit i dette sammendrag er hovedproblemet i forbindelse med økologisk inferens at specificere, under hvilke betingelser man kan slutte sig til individuel adfærd ud fra økologisk (aggregeret) adfærd. Den matematiske side af dette problem består i at formulere matematiske modeller for individuel adfærd og derefter udlede konsekvenserne af disse modeller for det aggregerede niveau, givet bestemte antagelser om forholdet mellem individuel og økologisk variation. For at kunne gennemføre økologisk inferens er det endvidere vigtigt, at det også er matematisk muligt "at gå den modsatte vej", dvs. fra det aggregerede til det individuelle niveau.

Selvfølgelig er det ikke nok at konstruere matematiske modeller, som i princippet tillader inferens imellem niveauer. Det er lige så vigtigt at tage i betragtning, om de adfærdsmodeller, der er indbygget i den matematiske teori, kan anvendes til beskrivelse af faktisk adfærd.

Da den lineære logit model for binært valg var så velegnet til økologisk analyse af faktisk vælgeradfærd, var det naturligt at søge efter en matematisk teori fra hvilken den økologiske lineære logit model kunne udledes. Desværre viste det sig, som nævnt i kapitel 3, afsnit 3.1, at hvis den multidimensionale lineære logit model blev antaget at gælde på det individuelle nivea, kunne dette ikke føre til udledning af en analog multidimensional lineær logit model på det økologiske niveau. Det gælder imidlertid, at hvis den multidimensionale lineære probit model, der er en meget tæt approksimation til logit modellen, istedet antages at gælde på det individuelle niveau, leder dette faktisk til udledningen af en analog probit model på det økologiske niveau (jvf. kaptitel 3, afsnit 3.3).

For at gøre økologisk inferens mulig med denne probit teori, er det nødvendigt at indføre bestemte antagelser om forholdet mellem den latente struktur på individniveau og den latente struktur på økologisk niveau. Under antagelse af isomorfi, dvs. at kovarians-matricen på det individuelle niveau er proportional med kovarians-matricen på det økologiske niveau; og den yderligere antagelse, at forholdet mellem individuel og økologisk varians antager en meget stor værdi, viser det sig, at den totale individuelle tetrakoriske korrelation er tilnærmet lig med den økologiske probit korrelation (jvf. kapitel 3, afsnit 3.5).

Disse to antagelser om henholdsvis isomorfi og høj varians rate formodes kun at være nogenlunde gyldige inden for homogene politiske regioner, hvor den økologiske variation mellem valgdistrikter er af samme art (har den samme dimensionale struktur) som variationen mellem individuelle vælgere (jvf. op. cit.). Da endvidere den økologiske logit korrelation er en tæt approksimation til den økologiske probit korrelation, er hypotesen bag den økologiske estimation af individuel vælgeradfærd:

For binært valg inden for en homogen politisk region er den totale individuelle tetrakoriske korrelation lig med den økologiske logit korrelation.

Dette resultat er noget forvirrende, for er der noget som fortalerne såvel som kritikerne af økologisk inferens er enige om, så er det, at den individuelle korrelation sjældent er lig med, men sædvanligvis er svagere end den økologiske korrelation(jvf. kapitel 3, afsnit 2.3.). Denne forvirring kan dog i nogen grad fjernes, hvis man gør sig klart, at det sædvanligvis er Pearson's produkt-moment korrelation, der beregnes på individniveau, selv i tilfælde hvor de korrelerede variable er binære. Det viser sig således, at det mest citerede eksempel på såkaldt "økologisk fejlslutning", oprindelig fremsat af den barske kritiker af økologisk inferens, Robinson, omhandlende korrelationen mellem race og analfabetisme, faktisk viser, at den individuelle korrelation er omtrent lig med den økologiske korrelation, hvis den tetrakoriske korrelation beregnes i stedet for produkt-moment korrelationen.

Økologisk inferens for multipelt valg

Som nævnt i forbindelse med diskussionen af økologiske modeller viste det sig i det danske valgprojekt, at modellen for binært valg er et lige så godt instrument til økologisk analyse af vælgeradfærd i et flerpartisystem som den mere komplekse model for multipelt valg. Dette gælder imidlertid ikke, når beskrivelsesniveauet er de individuelle vælgere. Bruger man således hypotesen fra det forrige afsnit til direkte økologisk estimation af individuel vælgeradfærd, viser det sig, at der er en klar tilbøjelighed til at overestimere vælgermobiliteten mellem partierne (kapitel 3, afsnit 4.1). Forklaringen på denne tilbøjelighed er, at partivalg opfattes som et "groft" binært valg imellem et bestemt parti og alle andre partier i stedet for som et "ægte" binært valg mellem to forskellige partier (afsnit 4.2).

Dette resultat var en alvorlig forhindring for anvendelse af teorien for binært valg på vælgeradfærd, for selv i toparti systemer er der også den tredje valgmulighed, at man afholder sig fra at stemme. Ved dette dødvande i forskningsprocessen viste den tidligere kasserede multiple logit model sin styrke. Selvom det ikke er muligt at bruge denne model til direkte økologisk estimation af individuel vælgeradfærd, viser modellen, at en indirekte fremgangsmåde er mulig ved simultan betragtning af forskellige binære delmængder af vælgerens valgmuligheder.

Problemet med at behandle mange forskellige binære delmængder simultant blev løst ved en computer iterations procedure. Med denne procedure anvendes den førnævnte hypotese for ægte binært valg ved simultan behandling af alle de mulige binære valg, der hver for sig involverer et bestemt "referenceparti" (for eksempel det "neutrale" sofavælgerparti) samt et af alle de andre partier. Ved denne iterations procedure konstruerer computeren gradvist alle mulige tetrakoriske (fire-felts) tabeller, som simultant opfylder hypotesen om identisk individuel og økologisk korrelation; og dermed er vælgermobiliteten mellem alle partier estimeret.

Selvom denne estimationsprocedure er udviklet på grundlag af en teori om vælgermobilitet mellem to valg, viser det sig, at nøjagtigt det samme computer program kan estimere sammenhængen mellem klasse og partivalg, hvis man erstatter de økologiske data om partivalg ved det ene valg med økologiske data om de forskellige socialklasser.

Økologisk estimation af vælgeradfærd i Danmark

Økologiske estimationer af vælgermobilitet og klassernes vælgeradfærd ved folketingsvalgene 1920-79 er vist i appendix 2-3. I kapitel 4 bliver validiteten af disse økologiske estimater vurderet ved sammenligning af økologisk estimeret vælgermobilitet mellem folketingsvalgene i 1971 og 1973 med to interview-panelundersøgelser samt ved sammenligning af økologiske estimater af klassernes vælgeradfærd i 1973 med to forskellige tværsnits interviewundersøgelser. På grundlag af disse sammenligninger konkluderes det, at validiteten af de økologiske estimater er på samme niveau som validiteten af interviewundersøgelserne.

Det er dog vigtigt at være opmærksom på, at de økologiske estimaters manglende validitet er tilbøjelig til at være forårsaget af systematiske fejl snarere end af tilfældige fejl, hvilket er i modsætning til, hvad der almindeligvis gælder for interviewundersøgelser. Denne tilbøjelighed til systematiske fejl kan bruges til de økologiske estimaters fordel i forbindelse med dynamisk analyse. Eftersom den systematiske skævhed (tilbøjeligheden til enten overestimation eller underestimation af procentandele) ser ud til at være temmelig stabil over tid, kan denne skævhed i et vist omfang "forkortes væk", når man studerer omfanget af forandring.

Som et eksempel kan nævnes, at omfanget af "bruttovandringer" mellem partier er tilbøjelig til at blive overestimeret inden for gruppen af landpartier samt inden for gruppen af bypartier, mens brutto-vandringer mellem disse to grupper er tilbøjelige til at blive underestimeret. Ser man imidlertid på parvise "nettovandringer" mellem partier, bliver en del af skævheden i estimationen tilsyneladende udjævnet.

Den samme tankegang kan anvendes, når man studerer forandringen i et partis eller en gruppe af partiers "klasseprofil". Til trods for at der kan være en skævhed i den økologisk estimerede tilslutning til et parti eller til en gruppe af partier, så er omfanget af forandring i denne tilslutning formodentlig mindre skævt estimeret.

Når dette "dynamiske princip" anvendes på tabellerne i appendix 2-3, ser resultaterne ud til at være temmelig valide. Ikke mindst de parvise nettovandringer ser ud til at være i overensstemmelse med, hvad man skulle forvente ud fra den danske politiske historieskrivning.

I afsnit 3 i kapitel 4 anvendes et index for klassepolarisering, udledt fra den økologiske analyse model, på den økologisk estimerede tilslutning til arbejderpartierne inden for henholdsvis arbejderklassen og middelklassen. Det er måske et indirekte bevis på de økologiske estimaters dynamiske validitet at man med dette index finder en meget klar sammenhæng mellem regeringens politiske sammensætning og forandringen i klassepolarisering i de enkelte valgperioder. Det ser således ud til, at det i perioden 1920 til 1979 har været en nødvendig (men ikke tilstrækkelig) forudsætning for faldende klassepolarisering i en valgperiode, at regeringen var socialdemokratisk ledet.

Videreudvikling af metoden til økologisk inferens

Efter min mening er den relativt høje validitet ved de resultater man opnår med den omtalte metode til økologisk inferens et tegn på, at latent struktur tilgangen er bedre end den for tiden fremherskende regressions tilgang. Dette gælder i særlig grad, når man tager i betragtning, at regressions tilgangen giver ugyldige resultater, hvis den økologiske estimation af individuel vælgeradfærd involverer mere end to partier eller mere end to klasser (kapitel 3, afsnit 2). Til forskel herfra frembyder antallet af partier eller klasser ikke noget problem for den metodologi, der baserer sig på latent struktur tilgangen.

Selvfølgelig er den metode, der omtales i denne bog, ikke det endelige svar på, hvordan man kan anvende latent struktur tilgangen til økologisk inferens inden for valgforskningen. Det er faktisk således, at de to udvalgte antagelser inden for latent struktur logit-teorien, nemlig antagelsen om isomorfi og antagelsen om meget høj varians rate mellem individuelt og økologisk niveau, udgør temmeligt ekstreme forenklinger. Det ser således ud til, at antagelsen om isomorfi ikke er helt valid, når man tager i betragtning, at land-by dimensionen er tilbøjelig til at influere de økologiske estimater for meget i relation til middelklasse-arbejderklasse dimensionen.

Dette resultat tyder på, at man kan opnå mere valide estimater med latent struktur teorien ved at formulere mere raffinerede antagelser om forholdet mellem det individuelle og det økologiske niveau - en opgave for den fremtidige udvikling af metoden til økologisk inferens.