

Temporal Dependence in Longitudinal Paired Comparisons

Dittrich, Regina; Francis, Brian; Katzenbeisser, Walter

DOI:

[10.57938/0c5cd849-18bf-4116-a360-c26220c9a7fc](https://doi.org/10.57938/0c5cd849-18bf-4116-a360-c26220c9a7fc)

Published: 01/10/2008

Document Version

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Dittrich, R., Francis, B., & Katzenbeisser, W. (2008). *Temporal Dependence in Longitudinal Paired Comparisons*. Research Report Series / Department of Statistics and Mathematics No. 67
<https://doi.org/10.57938/0c5cd849-18bf-4116-a360-c26220c9a7fc>

Temporal Dependence in Longitudinal Paired Comparisons



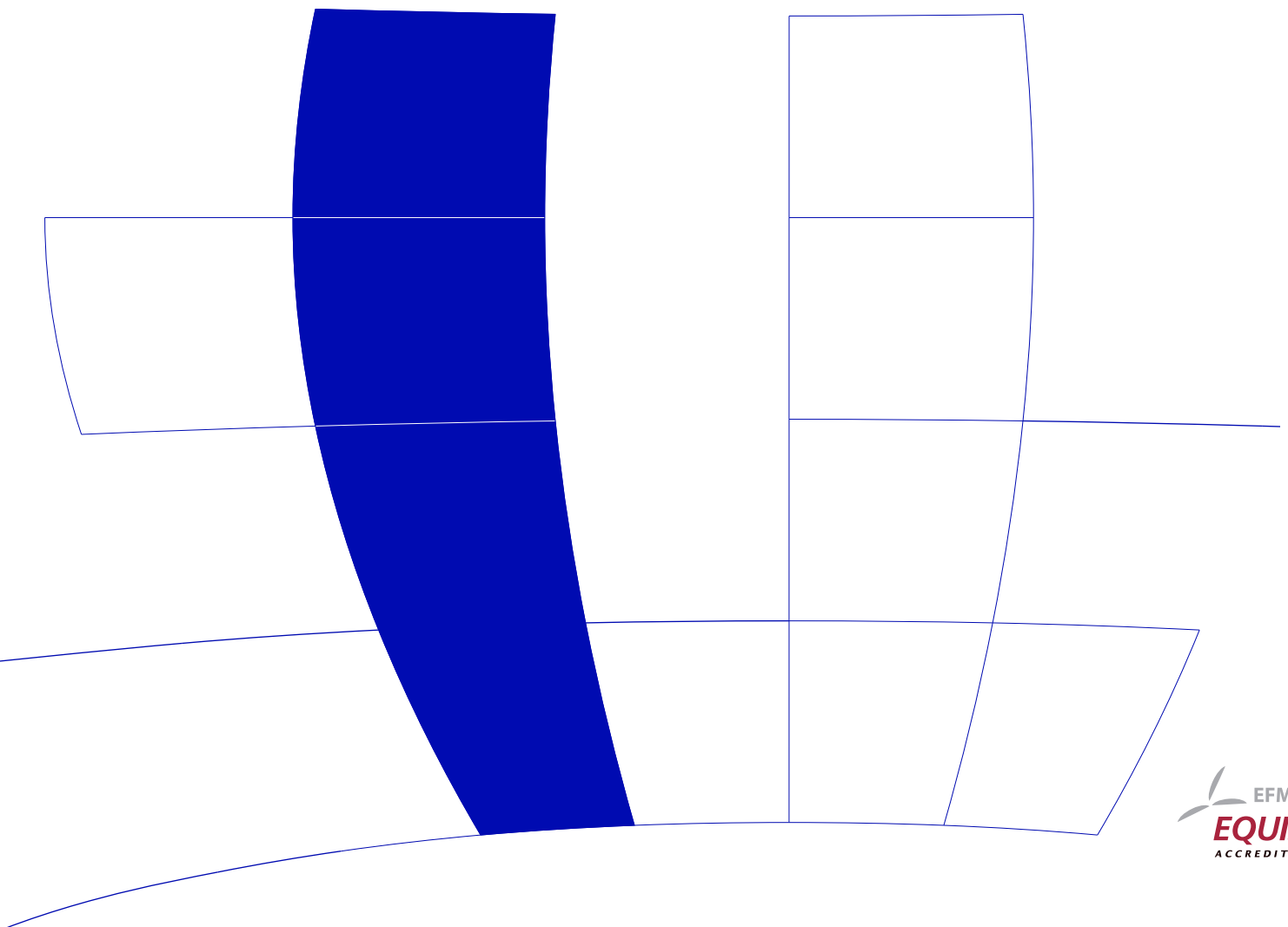
Regina Dittrich, Brian Francis, Walter Katzenbeisser

Department of Statistics and Mathematics
Wirtschaftsuniversität Wien

Research Report Series

Report 67
July 2008

<http://statmath.wu-wien.ac.at/>



Temporal Dependence in Longitudinal Paired Comparisons

Regina Dittrich
Wirtschaftsuniversität Wien

Brian Francis
University of Lancaster

Walter Katzenbeisser
Wirtschaftsuniversität Wien

Abstract

This paper develops a new approach to the analysis of longitudinal paired comparison data, where comparisons of the same objects by the same judges are made on more than one occasion. As an alternative to other recent approaches to such data, which are based on Kalman filtering, our approach treats the problem as one of multivariate multinomial data, allowing dependence terms between comparisons over time to be incorporated. The resulting model can be fitted as a Poisson log-linear model and has parallels with the quadratic binary exponential distribution of Cox. An example from the British Household Panel Survey illustrates the approach.

Keywords: Bradley-Terry model, Multiple multinomial responses, Log-linear model, Conditional log-odds ratio, Paired comparison data

1 Introduction

A common method of addressing the problem of ordering a set of objects on a preference scale is to utilise the method of paired comparisons. A paired comparison experiment will present two of a set of J objects to an individual and determine which of the two is preferred, either overall, or on some dimension such as style or colour. The experiment is then repeated, with a selection of pairs of objects being presented to a collection of individuals. For small J it is common to present all possible pairs to the respondent. This basic paired comparison approach has also been used to produce rankings of sports teams (where the judgement is produced by the final score of a game) and in analysing ranking data (Dittrich et al., 2000).

Most analyses of paired comparison data has used the Bradley-Terry model (Bradley and Terry, 1952) as a starting point. This model represents the unknown preference values or 'worths' of the objects as a set of non-negative values which lie between zero and one. The model can easily be fitted as a log-linear model. Since Bradley and Terry's original work, various authors have proposed additions to the basic model. Some of these relate to extensions of the paired comparison experimental design. For example, ties can be allowed as a preference choice (Rao and Kupper, 1967; Davidson, 1970; Kousgaard, 1976), and degrees of preference can be requested, leading to the development of ordinal

paired comparison models (Tutz, 1986; Agresti, 1992; Böckenholt and Dillon, 1997). Additionally, preferences can be requested on more than one attribute of an object (such as price, style and functionality), and this has led to the development of the multivariate paired comparison model (Davidson and Bradley, 1969; Böckenholt, 1988). Other developments relate solely to statistical modelling extensions, such as the incorporation of order effects (Davidson and Beaver, 1977; Fienberg, 1979) and the inclusion of explanatory variables on either or both of the subjects or objects as linear (Kousgaard, 1984; Matthews and Morris, 1995; Dittrich et al., 1998) or nonlinear effects (Francis et al., 2002). Models incorporating non-temporal dependence between objects have recently been proposed by Dittrich et al. (2002).

In this article, we are concerned with *time* dependence in the context of *longitudinal* paired comparison data, where repeated judgements are made on the same pairs of objects by the same judges. Such data arise in various circumstances, and temporal changes in preference can be present for different reasons:

- *changes in item characteristics over time.* A typical source of data would be a sports competition, where teams or individual players meet repeatedly over time. Here the worths of the athletes or teams (the objects) represent abilities or skills. These abilities may well change over time, but the judgements are given by the results of competitions and the judgement mechanism remain fixed if referee bias is ignored.
- *Changes in judgements over time.* An alternative scenario is for the objects to stay unchanged over time, and for the judgements to change. For example, an individual's preferences in music may well change as he or she matures.
- *Changes in object characteristics and judgements over time.* Both types of temporal shift can occur together - if the objects are political parties, for example, then the parties can change their policies while the judges (the voters) are changing their views on the worths of the parties.

If there is no time dependence, then models for longitudinal data can be fitted easily by estimating the worth parameters separately for each time point. However, it should be expected that temporal dependence is present, and a respondent who ranks two objects one way at one time point is more likely to rank those objects the same way at the next time point.

Both Fahrmeier and Tutz (1994) and Glickman (1999) have extended the usual (static) Bradley-Terry model by including parameters which allow for time dependence. Fahrmeier and Tutz introduced dynamic stochastic models for time-dependent ordered paired comparisons. The general model is characterized by the response or observation model, which connects observations and underlying abilities, and by the transitions model, which specifies the variation of the underlying abilities and parameters over time. These two components together form a non-Gaussian state-space model. The estimation method is based on an extension of Kalman filtering and smoothing for dynamic generalized linear models. Glickman (2001) extended the state-space models for paired comparisons by allowing not only the merits of the objects but also the variance of the state process to change stochastically. These methods are however inadequate

when the number of objects to be compared is large, because the computations involved become intractable. A computationally less intensive method for fitting large dynamic paired comparison models is developed by Glickman (1999).

In contrast to the above approach, the purpose of this paper is to discuss a simple log-linear representation for the dynamic (time-dependent) paired comparison Bradley-Terry model. The proposed method represents the observed paired comparisons as a set of multiple multinomial responses; this allows a multiplicative specification for the underlying probability distribution to be used, which has parallels to the quadratic exponential binary distribution (Cox, 1972; Cox and Wermuth, 1994). This framework can easily incorporate parameters which represent time dependencies. The advantage of this specification is that model fitting and model checking can easily be done within the Generalised Linear Modelling (GLM) framework. Moreover the parameters associated with the time dependencies can be interpreted as log-odds and as log-odds ratios.

2 A log-linear representation for longitudinal paired comparisons

We consider the situation where where J objects are compared in a paired comparison experiment repeatedly over time points t , $t = 1, 2, \dots, T$, by N judges, and where it is assumed that each judge responds at each time point to all paired comparisons. Furthermore it is assumed that the decisions between different judges are independent. At each time point t , this experiment results in $\binom{J}{2}$ paired comparisons, say in the pre-defined order

$$(1, 2), (1, 3), \dots, (1, J); (2, 3), \dots, (2, J); \dots; (J-1, J),$$

where (i, j) is a shorthand notation for the comparison of objects O_i and O_j . For the comparison of objects O_i and O_j at time point t there are three possible responses: (i) *preference for object O_j* , (ii) *no preference*, and (iii) *preference for object O_i* . These responses can be interpreted as realisations of the random variable Y_{ijt} defined by

$$Y_{ijt} = \begin{cases} -1 & \text{if object } O_j \text{ is preferred over } O_i \text{ at time } t, \\ 0 & \text{if there is no preference between } O_i \text{ and } O_j \text{ at time } t, \\ 1 & \text{if object } O_i \text{ is preferred over } O_j \text{ at time } t. \end{cases}$$

For any judge, the response pattern vector \mathbf{y} which represents the responses for all paired comparisons and all time points can be written as

$$\mathbf{y} = (y_{121}, y_{122}, \dots, y_{12T}; y_{131}, y_{132}, \dots, y_{13T}; \dots; y_{J-1, J1}, y_{J-1, J2}, \dots, y_{J-1, JT}),$$

a vector of $T \binom{J}{2}$ elements each consisting of one of the values $\{-1, 0, 1\}$, and in the standard order of a $3^{T \binom{J}{2}}$ factorial main effects only design; a few of these

response pattern vectors are then given by

$$\begin{aligned}
\mathbf{y}_1 &= (-1, -1, \dots, -1, -1), \\
\mathbf{y}_2 &= (-1, -1, \dots, -1, 0), \\
\mathbf{y}_3 &= (-1, -1, \dots, -1, 1), \\
\mathbf{y}_4 &= (-1, -1, \dots, 0, -1), \\
\mathbf{y}_5 &= (-1, -1, \dots, 0, 0), \\
\mathbf{y}_6 &= (-1, -1, \dots, 0, 1), \\
&\vdots \\
\mathbf{y}_L &= (1, 1, \dots, 1, 1).
\end{aligned}$$

where $L = 3^{T\binom{J}{2}}$ gives the total number of possible response pattern vectors. Thus each judge will give a response which is one of the L response pattern vectors.

Following a similar approach to the multivariate paired comparison model (Böckenholt, 1988), it is assumed that there exists an underlying latent scale on which the parameters $\pi_{1t}, \pi_{2t}, \dots, \pi_{Jt}$ representing the 'worth' of the objects at each time t are located. Identifiability is as usual achieved by setting $\sum_{i=1}^J \pi_{it} = 1$. As we are allowing for a 'no preference' response (a tied judgement), the basic Bradley-Terry model is not appropriate. Instead we take as a starting point the *Adjacent Categories* model (Böckenholt and Dillon (1997)) which was originally proposed for modelling ordinal paired comparison experiments, by postulating a power relationship between the response categories and the probability of preferring object O_j over object O_i . We show below that a basic paired comparison experiment involving ties can be modelled as a special case of an ordered paired comparison model.

Let $\{Y_{ijt} = y_{ijt}\}, y_{ijt} \in \{-1, 0, 1\}$, the adjacent categories model can be written as

$$P\{Y_{ijt} = y_{ijt} | \pi_{it}, \pi_{jt}, \nu_{ht}\} = a_{ijt} \nu_{ht} \left(\frac{\pi_{it}}{\pi_{it} + \pi_{jt}} \right)^{1+y_{ijt}} \left(\frac{\pi_{jt}}{\pi_{it} + \pi_{jt}} \right)^{1-y_{ijt}}, \quad (1)$$

where all parameters are positive, and

$$\nu_{ht} = \begin{cases} 1 & , \text{ if } y_{ijt} = \pm 1, \\ \nu_{0t} & , \text{ if } y_{ijt} = 0. \end{cases}$$

a_{ijt} denotes a normalizing constant in order to make the probabilities in (1) sum to unity. ν_{0t} can be interpreted as a parameter representing *no decision*. In order to obtain a log-linear representation we follow Sinclair (1982) in rewriting these probabilities as follows:

$$\frac{\pi_{it}}{\pi_{it} + \pi_{jt}} = \frac{\sqrt{\pi_{it}/\pi_{jt}}}{\sqrt{\pi_{it}/\pi_{jt}} + \sqrt{\pi_{jt}/\pi_{it}}}, \quad \frac{\pi_{jt}}{\pi_{it} + \pi_{jt}} = \frac{\sqrt{\pi_{jt}/\pi_{it}}}{\sqrt{\pi_{it}/\pi_{jt}} + \sqrt{\pi_{jt}/\pi_{it}}}. \quad (2)$$

This produces the following representation

$$P\{Y_{ijt} = y_{ijt} | \pi_{it}, \pi_{jt}, \nu_{ht}\} = a_{ijt}^* \nu_{ht} \left(\frac{\sqrt{\pi_{it}}}{\sqrt{\pi_{jt}}} \right)^{1+y_{ijt}} \left(\frac{\sqrt{\pi_{jt}}}{\sqrt{\pi_{it}}} \right)^{1-y_{ijt}}, \quad (3)$$

where $a_{ijt}^* = \frac{a_{ijt}}{(\sqrt{\pi_{it}/\pi_{jt}} + \sqrt{\pi_{jt}/\pi_{it}})^2}$ is another normalizing constant which will be treated as nuisance. Finally, equation (3) can conveniently be rewritten as

$$P\{Y_{ijt} = y_{ijt}\} = a_{ijt}^* \left(\frac{\pi_{it}}{\pi_{jt}}\right)^{y_{ijt}} (\nu_{0t})^{1-|y_{ijt}|}, \quad y_{ijt} \in \{-1, 0, 1\}. \quad (4)$$

Equation (4) is the basic building block for further development and can be seen to be the Davidson model for ties in paired comparison experiments ((Davidson, 1970)).

Assuming independent decisions the joint distribution of the response vectors is given by

$$P\{\mathbf{Y} = \mathbf{y}\} = \Delta \prod_{i < j} \prod_t \left(\frac{\pi_{it}}{\pi_{jt}}\right)^{y_{ijt}} (\nu_{0t})^{1-|y_{ijt}|}, \quad \mathbf{y}_{ijt} \in \{-1, 0, 1\}, \quad (5)$$

where Δ is another normalising constant.

2.1 Model Specification

In section 1, we have already discussed possible mechanisms which point to the need to allow for temporal dependence in a statistical model. We proceed by making two assumptions:

- (i) Dependencies are introduced solely by repeated evaluation of the same object pair (O_i, O_j) by the same judge over different time points $t = 1, 2, \dots, T$. Thus, we are primarily concerned with assessment of temporal dependence and therefore assume for simplicity that decisions concerning different object pairs are independent. However, this assumption can be relaxed and non-temporal dependencies easily introduced; the discussion has further details.
- (ii) A Markovian structure is assumed, which means that only the previous decision (at time $t - 1$) has an influence on the decision at time t for the given comparison of objects i and j , and which will be represented by additional parameters $\theta_{ij|t-1,t}$.

In order to obtain a simple log-linear model that accounts for time dependencies between the judgments, we start by specifying the joint distribution of the random variables

$$(Y_{121}, \dots, Y_{12T}; Y_{131}, \dots, Y_{13T}; \dots; Y_{J-1,J1}, \dots, Y_{J-1,JT}) = \mathbf{Y}.$$

according to both assumptions in a multiplicative way: The joint distribution of \mathbf{Y} will be specified by

$$P\{\mathbf{Y} = \mathbf{y}\} = \Delta^* \prod_{i < j} \left(\frac{\pi_{i1}}{\pi_{j1}}\right)^{y_{ij1}} (\nu_{01})^{1-|y_{ij1}|} \prod_{t=2}^T \left(\frac{\pi_{it}}{\pi_{jt}}\right)^{y_{ijt}} (\nu_{0t})^{1-|y_{ijt}|} \exp\{\theta_{ij|t-1,t} y_{ij,t-1} y_{ijt}\} \quad (6)$$

where $y_{ijt} \in \{-1, 0, 1\}$, and Δ^* is a further normalising constant. Therefore we include temporal dependence in the model by including terms of the form

$$\exp\{\theta_{ij|t-1,t} y_{ij,t-1} y_{ijt}\},$$

because for $\theta_{ij|t-1,t} = 0$ we get the independence model (5). The inclusion of terms of this form is similar to the approach taken by Cox (Cox, 1972; Cox and Wermuth, 1994) when modelling dependence in multivariate binary data, who termed the resulting distribution the quadratic exponential binary distribution. The advantage of this specification is that the parameters $\theta_{ij|t-1,t}$ can easily be interpreted as log-odds ratios as will be shown in section 2.3; moreover temporal dependencies of higher order can easily be incorporated in the model.

2.2 Estimation of the Parameters

Let $p(\mathbf{y}_\ell)$ denote the probability that response pattern vector \mathbf{y}_ℓ occurred. Then (6) can be used to calculate these probabilities. In addition, let N_ℓ denote the random variable

$N_\ell =$ number of times where response pattern vector \mathbf{y}_ℓ , $\ell = 1, 2, \dots, L$ occurs,

then the N_ℓ 's are multinomially distributed with $N = \sum_\ell N_\ell$ and with probabilities $p(\mathbf{y}_\ell)$. The expectations m_ℓ of N_ℓ are given by $m_\ell = Np(\mathbf{y}_\ell)$ and can be represented in a log-linear way, therefore we base our parameter estimation on simple multinomial sampling.

Because model (6) belongs to the class of generalized linear models the parameters can be estimated by standard software, for example by GLIM (Francis et al., 1993), using a Poisson distribution and a log-link. The design matrix consists of column vectors with suitable entries for the parameters δ , λ , γ , and θ .

In the general case the design matrix for model (6) is given by

$$\mathbf{X} = (\mathbf{1}_L, \mathbf{YA}, \mathbf{C}, \mathbf{W}) ,$$

where $\mathbf{1}_L = (1, 1, \dots, 1)$ is a column vector of length $L = 3^{T \binom{J}{2}}$ elements representing the parameter δ . The $(L \times T \binom{J}{2})$ matrix \mathbf{Y} , the response pattern matrix, is the design matrix for a $3^{T \binom{J}{2}}$ main effects only design in standard order (in fact the rows of \mathbf{Y} are the response pattern vectors \mathbf{y}_ℓ) and

$$\mathbf{A} = \mathbf{B} \otimes \mathbf{I}_T .$$

The $(\binom{J}{2} \times J)$ matrix \mathbf{B} is the so called paired comparison design matrix (Böckenholt and Dillon (1997)). Each column of this matrix corresponds to one of the J objects, and each row to one of the $\binom{J}{2}$ paired comparisons (in the given order) and is given by

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix} .$$

\mathbf{I}_T is the corresponding identity matrix of order T .

The matrix \mathbf{C} can be written as

$$\mathbf{C} = \begin{pmatrix} J \\ 2 \end{pmatrix} \cdot \mathbf{J} - (\mathbf{Y} \odot \mathbf{Y}) \cdot (\mathbf{1}_{\binom{J}{2}} \otimes \mathbf{I}_T) ,$$

where \mathbf{J} is a $(L \times T)$ matrix consisting of ones, and \odot denotes the Hadamard product of the response pattern matrix, i.e. the elementwise product of the matrix \mathbf{Y} with itself.

The matrix \mathbf{W} represents two-way interactions between paired comparisons evaluated over adjacent time points. The columns of \mathbf{W} can be generated in the usual way as two-way or higher-way interactions, i.e. by elementwise multiplication of corresponding columns of \mathbf{Y} . For more than two time points there could be various dependency structures: for example one could either consider a Markovian-type structure as before, i.e. an autoregressive structure of order 1 (which means that one has to consider interactions of the type $y_{122}y_{121}, y_{123}y_{122}, y_{124}y_{123}, \dots$) or structures incorporating higher order dependencies into the Markovian model. A further advantage of this approach is that various types of dependencies can be tested in the usual way by comparing the deviances of the suitable nested models.

2.2.1 Example

For illustrative purposes we consider the simplest non-trivial case, $J = 3$ and $T = 2$. For the probability $P\{Y_{121} = y_{121}, Y_{122} = y_{122}; Y_{131} = y_{131}, Y_{132} = y_{132}; Y_{231} = y_{231}, Y_{232} = y_{232}\}$ with $y_{ijt} \in \{-1, 0, 1\}$ and $i, j = 1, 2, 3, i < j$ we get

$$\begin{aligned} P\{Y_{121} = y_{121}, Y_{122} = y_{122}; Y_{131} = y_{131}, Y_{132} = y_{132}; Y_{231} = y_{231}, Y_{232} = y_{232}\} = \\ = \Delta^* \prod_{i < j} \left(\frac{\pi_{i1}}{\pi_{j1}} \right)^{y_{ij1}} (\nu_{01})^{1-|y_{ij1}|} \left(\frac{\pi_{i2}}{\pi_{j2}} \right)^{y_{ij2}} (\nu_{02})^{1-|y_{ij2}|} \exp\{\theta_{ij|1,2} y_{ij1} y_{ij2}\} \end{aligned}$$

Therefore the probability of the first response pattern vector $\mathbf{y}_1 = (-1, -1, -1, -1, -1, -1)$ is for example given by:

$$\begin{aligned} p(\mathbf{y}_1) &= \Delta^* \frac{\pi_{21}}{\pi_{11}} \frac{\pi_{22}}{\pi_{12}} \frac{\pi_{31}}{\pi_{11}} \frac{\pi_{32}}{\pi_{12}} \frac{\pi_{31}}{\pi_{21}} \frac{\pi_{32}}{\pi_{22}} \times \exp\{\theta_{12|12} + \theta_{13|12} + \theta_{23|12}\} \\ &= \Delta^* \pi_{11}^{-2} \pi_{12}^{-2} \pi_{31}^2 \pi_{32}^2 \times \exp\{\theta_{12|12} + \theta_{13|12} + \theta_{23|12}\} \end{aligned}$$

and the logarithm of the expectation of N_1 is given by

$$\ln m_1 = \delta - 2\lambda_{11} - 2\lambda_{12} + 2\lambda_{31} + 2\lambda_{32} + \theta_{12|12} + \theta_{13|12} + \theta_{23|12}, \quad (7)$$

where $\lambda_{ij} = \ln \pi_{ij}$, $\gamma_{0t} = \ln \nu_{0t}$ and the nuisance parameter $\delta = \ln\{\Delta^* N\}$. For model (6) the following log-linear representation for the expectation of the number of times when response pattern vector \mathbf{y}_ℓ occurred can be written in matrix notation as

$$\begin{pmatrix} \ln m_1 \\ \ln m_2 \\ \ln m_3 \\ \ln m_4 \\ \ln m_5 \\ \ln m_6 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & -2 & -2 & 0 & 0 & 2 & 2 & 0 & 0 & 1 & 1 & 1 \\ 1 & -2 & -2 & 0 & 1 & 2 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & -2 & -2 & 0 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & -1 \\ 1 & -2 & -2 & 1 & 0 & 1 & 2 & 1 & 0 & 1 & 1 & 0 \\ 1 & -2 & -2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & -2 & -2 & 1 & 2 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \times \begin{pmatrix} \delta \\ \lambda_{11} \\ \lambda_{12} \\ \lambda_{21} \\ \lambda_{22} \\ \lambda_{31} \\ \lambda_{32} \\ \gamma_{01} \\ \gamma_{02} \\ \theta_{12|12} \\ \theta_{13|12} \\ \theta_{23|12} \end{pmatrix}$$

Because of the identifiability condition, the relationship between the initial λ -parameters and the final worth parameters π is given by

$$\pi_{it} = \frac{\exp\{\lambda_{it}\}}{\sum_{j=1}^J \exp\{\lambda_{jt}\}}. \quad (8)$$

Note that the coefficients associated with the parameters λ in (7) can be interpreted as *the number of times where object O_i is preferred at time t minus the number of times, where object O_i is not preferred at time t* for all comparisons involving object O_i . Similarly, the coefficient associated with the parameter indicating *no preference* can be interpreted as *the number of times when there were no decisions at time point t* .

2.3 Interpretation of the Parameters

One advantage of the specification of the model is that the marginal (joint) distribution of the random variables $\{Y_{ij1}, Y_{ij2}, \dots, Y_{ijT}\}$ has the particularly simple form given in (6), which allows the calculation of (conditional) log-odds to be carried out easily.

For example, the effect of the dependencies on the decisions of the judges is best seen by considering the conditional log-odds in favour of object O_i in the comparison with object O_j at the last time point T conditional on all previous decisions at time points $t = 1, 2, \dots, T - 1$:

$$\begin{aligned} \ln \frac{P\{Y_{ijT} = 1|Y^-\}}{P\{Y_{ijT} = -1|Y^-\}} &= \ln \frac{P\{Y_{ij1} = y_{ij1}, \dots, Y_{ij,T-1} = y_{ij,T-1}, Y_{ij,T} = 1\}}{P\{Y_{ij1} = y_{ij1}, \dots, Y_{ij,T-1} = y_{ij,T-1}, Y_{ij,T} = -1\}} = \\ &= 2(\lambda_{iT} - \lambda_{jT}) + 2\theta_{ij|T-1,T}y_{ij,T-1}, \end{aligned} \quad (9)$$

where $Y^- = (Y_{ij1}, \dots, Y_{ij,T-1})$. Therefore the log odds in favour of object O_i at time point T is not only dependent on the object parameters, as in the independence model, but there is also a possible carry over effect from the previous period. If $\theta_{ij|T-1,T}$ is positive there is an additional effect in favour of object O_i if $y_{ij,T-1} = 1$, i.e. there was also a decision in favour of object O_i in the previous period $T - 1$.

These log-odds are rather complicated when considering another time point $t < T$. Consider for example the marginal joint distribution of $Y_{ij1}, Y_{ij2}, \dots, Y_{ij,T-1}$ which is given by

$$\begin{aligned} P\{Y_{ij1} = y_{ij1}, \dots, Y_{ij,T-1} = y_{ij,T-1}\} &= \\ &= \Delta^* \left(\frac{\pi_{i1}}{\pi_{j1}}\right)^{y_{ij1}} (\nu_{01})^{1-|y_{ij1}|} \prod_{t=2}^{T-1} \left(\frac{\pi_{it}}{\pi_{jt}}\right)^{y_{ijt}} (\nu_{0t})^{1-|y_{ijt}|} \exp\{\theta_{ij|t-1,t}y_{ij,t-1}y_{ijt}\} \\ &\times \left[\frac{\pi_{jT}}{\pi_{iT}} \exp\{-\theta_{ij|T-1,T}y_{ij,T-1}\} + \frac{\pi_{iT}}{\pi_{jT}} \exp\{\theta_{ij|T-1,T}y_{ij,T-1}\} + \nu_{0T} \right] \times A, \end{aligned} \quad (10)$$

where A is a constant, which results from marginalising the joint distribution (6) with respect to all other comparisons except the comparisons including objects i and j . Because (10) is not a distribution of type (6) the underlying marginal

distribution of the random variables $Y_{ij1}, Y_{ij2}, \dots, Y_{ijT}$ is not closed under further marginalisation. However, in the special case when $\pi_{iT} = \pi_{jT}$, (10) can be written as

$$\begin{aligned} P\{Y_{ij1} = y_{ij1}, \dots, Y_{ijT-1} = y_{ijT-1}\} &= \\ &= \Delta^* \left(\frac{\pi_{i1}}{\pi_{j1}} \right)^{y_{ij1}} (\nu_{01})^{1-|y_{ij1}|} \prod_{t=2}^{T-1} \left(\frac{\pi_{it}}{\pi_{jt}} \right)^{y_{ijt}} (\nu_{0t})^{1-|y_{ijt}|} \exp\{\theta_{ij|t-1,t} y_{ij,t-1} y_{ijt}\} \\ &\times [2 \cosh\{\theta_{ij|T-1,T} y_{ij,T-1}\} + \nu_{0T}] \times A. \end{aligned}$$

Calculating the log-odds in the conditional distribution of $Y_{ij,T-1} | Y_{ij1}, \dots, Y_{ij,T-2}$ in a similar way to (9) we get

$$2(\lambda_{i,T-1} - \lambda_{j,T-1}) + 2\theta_{ij|T-2,T-1} y_{ij,T-2},$$

because $\cosh(x)$ is an even function. Moreover, it can be shown that this holds for all time points t , $t = 1, 2, \dots, T$ if it is assumed that $\pi_{i\tau} = \pi_{j\tau}$, $\tau > t$.

Another advantage of modelling temporal dependence in this way is that the parameters $\theta_{ij|t-1,t}$ representing the dependencies have a simple interpretation as log-odds ratios, however in the *conditional* distribution of two consecutive Y 's, i.e. $Y_{ij,t-1}$ and Y_{ijt} , where, because of the underlying Markovian structure, conditioning can be done with respect to the preceding variables $(Y_{ij1}, \dots, Y_{ij,T-2}) = Y^-$, which seems to be sensible in this time series context.

The conditional distribution of the random variables $Y_{ij,t-1}, Y_{ijt}$ given all previous Y 's can be displayed in a (3×3) table, where the entries $p_{m,n}$ are the conditional probabilities $P\{Y_{ij,t-1} = m, Y_{ijt} = n | Y^-\}$ with $m, n \in \{-1, 0, 1\}$.

		Y_t		
		-1	0	1
Y_{t-1}	-1	$p_{-1,-1}$	$p_{-1,0}$	$p_{-1,1}$
	0	$p_{0,-1}$	$p_{0,0}$	$p_{0,1}$
	1	$p_{1,-1}$	$p_{1,0}$	$p_{1,1}$

The dependency of these random variables can be described by nine odds ratios, formed by combining each of the three pairs of rows with each of the three pairs of columns in turn. However, these odds ratios can be generated by a minimal set of odds ratios. Taking odds ratios relative to the bottom right hand cell ($m = 1, n = 1$) there are four possible odds ratios forming such a minimal set (Agesti (1990), p.18):

$$\alpha_{-1,-1} = \frac{p_{-1,-1} p_{1,1}}{p_{1,-1} p_{-1,1}}, \alpha_{-1,0} = \frac{p_{-1,0} p_{1,1}}{p_{1,0} p_{-1,1}}, \alpha_{0,-1} = \frac{p_{0,-1} p_{1,1}}{p_{0,1} p_{1,-1}}, \alpha_{0,0} = \frac{p_{0,0} p_{1,1}}{p_{0,1} p_{1,0}}$$

where the subscripts m' and n' of α represent the indices of the top left cell of the 2×2 sub-table whereas the bottom right hand cell is fixed with the indices $(1, 1)$. It can be shown that these odds-ratios are all multiples of $\theta_{ij|t-1,t}$: with the multiplying factor defined by $(m - m') \times (n - n')$. For example, the odds-ratios in the conditional distribution of the random variables $Y_{ij,t-1}$ and Y_{ijt} given Y^- can be shown to be

$$\alpha_{-1,-1} = \exp\{4\theta_{ij|t-1,t}\}, \alpha_{-1,0} = \alpha_{0,-1} = \exp\{2\theta_{ij|t-1,t}\}, \alpha_{0,0} = \exp\{\theta_{ij|t-1,t}\}. \quad (11)$$

I'm going to read out some things that may concern you. I'd like you to give me the answer off this card that comes closest to how concerned you are about each of the following.					
		A great deal	A fair amount	Not very much	Not at all
a)	The rising price of food and other consumer goods	1	2	3	4
b)	The destruction of the ozone layer	1	2	3	4
c)	The high rate of unemployment	1	2	3	4
d)	The extinction of many animal and plant species	1	2	3	4
e)	Declining moral standards	1	2	3	4

Figure 1: Example: "Concerns of British households" question from the BHPS

It can also be shown that all 'local' log odds ratios, which are formed by using cells in adjacent rows and adjacent columns, are given by $\theta_{ij|t-1,t}$. Therefore a positive parameter $\theta_{ij|t-1,t}$ indicates a uniformly positive 'local' association between the random variables $Y_{ij,t-1}$ and $Y_{ij,t}$.

3 An Application to the British Household Panel Study

To illustrate this method, we take as an example a set of social attitude questions from the British Household Panel Study (BHPS) (Buck et al., 1994). The BHPS is an household-based survey, taking as a base 8,167 selected households in England, Wales and Scotland (excluding the far north of Scotland). All adult members of the 5,538 households which agreed to participate made up a panel of individuals. The initial wave of the survey took place in 1991, with over 10,000 individual panel members being surveyed. These individuals have been followed through time in annual successive waves of the survey. If these individuals joined new households, then these new household members have also been invited to join the panel. A feature of the BHPS is that questions asked in early waves are repeated in subsequent waves, allowing social change to be assessed. Thus longitudinal information is available over the ten annual waves of the study. The study is ongoing, and additional waves are planned.

We concentrated on the responses to a question which measures concern about various social and political issues of contemporary relevance. The complete question is displayed in Figure 1. This question has so far been administered three times - in sweeps 2, 4 and 6, which took place in 1992, 1994 and 1996.

This question consists of a series of four-point Likert scales which give the absolute level of concern for each of the five items. For illustrating this example, we simplified the analysis by considering only three items. We chose concerns

OZ b) *destruction of the ozone layer*, UN c) *rate of unemployment*, and MO e) *declining moral standards*, as they represented distinct concepts. The omitted items a) and d) duplicated to some extent the chosen concerns with a) and c) both concerned about economic issues, and items b) and d) both concerned with green issues. We then converted the data to paired comparison form by comparing the responses in pairs first OZ to UN, then OZ to MO and finally UN to MO. For each paired comparison, the values of the Likert scale response given were used to determine whether there was more concern for the first item of the pair, the second item of the pair, or whether the items were judged to be of equal concern: The first item was taken to be preferred to the second item within each pair if the Likert scale response of the first item was smaller than the Likert scale response of the second item, two items were considered to be equal if their Likert scale responses had the same value.

From over 30,000 individuals participating in the BHPS in the ten sweeps, we selected out the 7,179 respondents who were interviewed in sweep 2, sweep 4 and sweep 6. However, these respondents included many groups of individuals who belonged to the same household, and there is likely to be dependence between such individuals. We therefore randomly selected one individual from each household at sweep 2, and followed that individual into sweeps 4 and 6. We also ensured that any new households formed at later sweeps by two sample individuals forming a new household also only contributed one individual to the analysis. This reduced our sample to 4250 individuals. Finally, we removed all cases who did not give complete responses to the 'concerns' question on all three sweeps. This further reduced our sample to the final analysis sample of 4155 cases.

The design of our problem leads to 19,683 possible response patterns. Of these patterns, the most frequent is the response MO \succ UN \succ OZ at all three time points (425 respondents). 19,001 response patterns do not occur in the data.

The basic model with no temporal dependence (5) gives a scaled deviance of 29588 on 19673 degrees of freedom.

A remarkable improvement of the fit occurs when we incorporate a Markovian time dependency structure (6), with a decrease in scaled deviance of 5038 on 6 degrees of freedom. The parameter estimates and worth parameters from this model are shown in Table 1.

The changes in the worth parameters over time are of primary interest, and are illustrated in Figure 2. As can be seen, the relative concerns about the environmental, economic and moral issues have changed between 1992 and 1996 in different ways. The concern about the green issue measured by the concern about the destruction of the ozone layer remains relatively stable over the observed time period. The economic issue represented by the concern about the high unemployment rate is decreasing over the observed time period. This effect might be due to the decreasing levels of unemployment over the study period - the seasonally adjusted unemployment rate for the UK collected from the Labour Force Survey and available from the Office for National Statistics (1999) are 9.8% for 1992, 9.8% for 1998 and 8.3% for 1996. The increasing importance of the ethical issue suggests that people become more and more concerned that the moral standards are declining. One reason for this latter effect might be the increasing perception of lack of financial and moral standards of the government of John Major (1992-1997), although it is also important to

Table 1: Model with Markovian time dependency structure

estimates	s.e.	parameters	worth
-7.243	0.083	<i>DELTA</i>	
0.251	0.022	<i>OZ</i> _{TIME1}	0.291
0.761	0.023	<i>UN</i> _{TIME1}	0.484
0.000	aliased	<i>MO</i> _{TIME1}	0.226
-0.279	0.023	<i>OZ</i> _{TIME2}	0.273
0.013	0.023	<i>UN</i> _{TIME2}	0.366
0.000	aliased	<i>MO</i> _{TIME2}	0.361
-0.347	0.021	<i>OZ</i> _{TIME3}	0.281
-0.212	0.021	<i>UN</i> _{TIME3}	0.322
0.000	aliased	<i>MO</i> _{TIME3}	0.397
1.149	0.022	<i>GAMMA</i> _{TIME1}	
1.167	0.023	<i>GAMMA</i> _{TIME2}	
1.060	0.021	<i>GAMMA</i> _{TIME3}	
1.034	0.038	<i>THETA</i> _{OZ-UN TIME1-2}	
1.097	0.037	<i>THETA</i> _{OZ-UN TIME2-3}	
1.257	0.037	<i>THETA</i> _{OZ-MO TIME1-2}	
1.228	0.036	<i>THETA</i> _{OZ-MO TIME2-3}	
0.956	0.038	<i>THETA</i> _{UN-MO TIME1-2}	
1.017	0.037	<i>THETA</i> _{UN-MO TIME2-3}	

realise that with longitudinal data of this type, the increase may be also be due to an age effect, with an increase in moral concern as the participants age.

The *theta*-parameters can be interpreted by using formula (11). The odds ratios $\alpha_{-1,-1} = \frac{p_{-1,-1}p_{1,1}}{p_{1,-1}p_{-1,1}}$ can be interpreted as the chances for the *same* preferences $p_{-1,-1}$ or $p_{1,1}$ at two consecutive time points compared to a *switching response*- a change of two units $p_{1,-1}$ or $p_{-1,1}$. For example the odds ratios $\alpha_{-1,-1}$ for OZ and UN between TIME 1 and 2 is given by $\exp(4 * THETA_{OZ-UN|TIME1-2}) = \exp(4 * 1.034) = 62.6$ and between TIME 2 and 3 it is 80.5 . This means that the chances for showing a stable reaction at consecutive time points, either marking OZ or UN higher than the other at both time points, is much higher compared to a switching response, marking OZ higher than UN at time 1 but UN higher than OZ at time 2 or vice versa.

These switching odds ratios for the three comparisons (UN-MO),(OZ-UN) and (OZ-MO) vary in magnitude between 45.8 ($\exp(4*0.956)$) and 152.6 ($\exp(4*1.257)$) indicating a strong tendency towards stable reactions for all items over all time periods.

4 Discussion

The purpose of this paper was to introduce a log-linear model for the analysis of dynamic (time-dependent) paired comparisons. The model is useful as it can be fitted with standard software and in the well-known GLM framework. Thus model checking can be carried out and models selected using standard methods based on the likelihood.

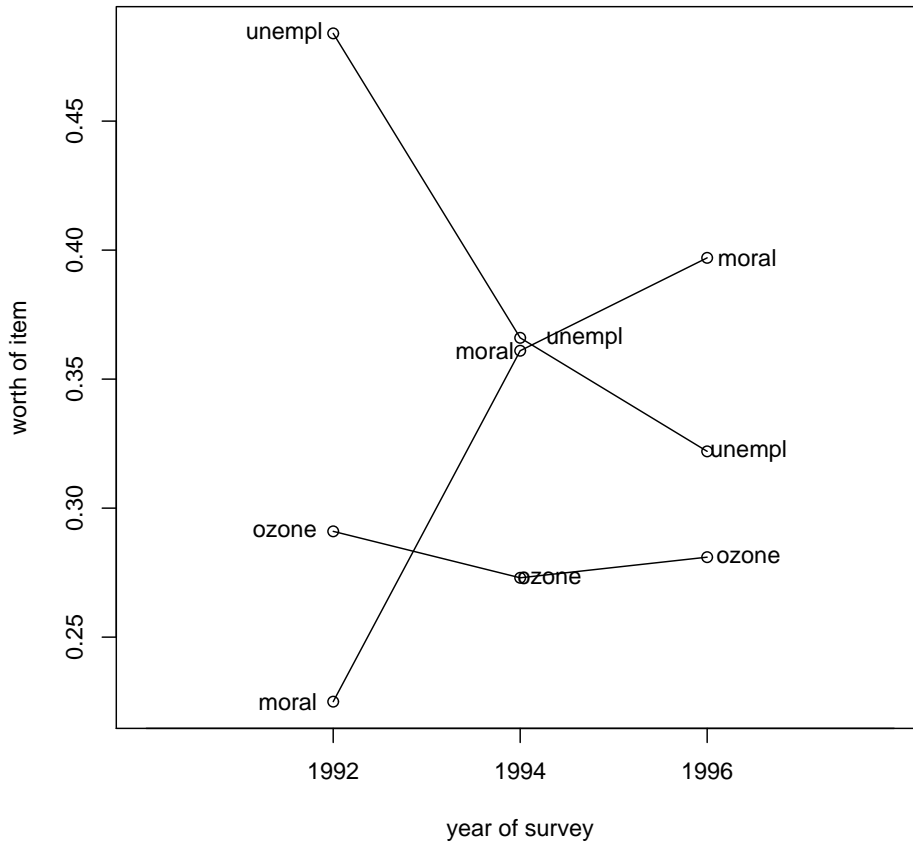


Figure 2: Changes in item worth parameters over time

One common requirement is to simplify the fitted model by imposing various restrictions and equality constraints in order to find a more parsimonious model. An advantage of this GLM approach is that all hypotheses can be tested in the usual way by examining deviance changes of the involved nested models.

For example, in the above illustration, we note that the three gamma parameters are very similar, and could be combined into a single common gamma parameter for all time points. Fitting this simplified model gives a deviance reduction of 14.75 on 2 degrees of freedom; there is strong evidence that we cannot simplify the model in this way. Similarly, it is not possible to simplify the model by setting all *thetas* to be equal, with a change in scaled deviance of 74.63 on 5 degrees of freedom. However, for each pair of items i and j , the θ_{ijs} for TIME 1-2 and TIME 2-3 can be equated, with a change in scaled deviance of 2.80 on 3 degrees of freedom. This model gives a uniformly positive local association over time, but with different association parameters for the three comparisons: $THETA_{UN-MO} = 0.986$ is the least important parameter

$THETA_{OZ-UN} = 1.067$ and $THETA_{OZ-MO} = 1.242$ is the most important parameter.

The log-linear approach to temporal dependence also allows the model to be extended in various ways.

Firstly, we could consider any potential ordinality on the paired comparison response. The underlying adjacent categories model (1) is designed to deal with such ordinality and can be easily extended to more than three response categories to incorporate such responses. For example, we could make more use of the Likert scale by measuring the *degree* to which one item is preferred to another. Items separated by only one Likert scale point could be rated as 'slightly preferred', items separated by two points could be rated as 'preferred' and so on. It is also possible to incorporate additional category parameters c , which would represent a tendency towards a particular response category independent of the items.

Simple covariate models can also be considered. For example, examining the effect of a single factor with H levels will mean that H distinct response vectors $\mathbf{y}_h, h = 1 \dots H$ (one for each level of the factor) will need to be formed and combined to form a new response vector $(\mathbf{y}_1, \mathbf{y}_2 \dots, \mathbf{y}_H)$. Examining the effect of gender in the above example would thus double the length of the response vector.

It is also possible to develop more complex dependencies. Firstly it is straightforward to incorporate higher order temporal dependencies. For example a second order autoregressive process can be taken into account, by augmenting the exp-term in (6) by $\theta_{ij|t-2,t}y_{ij,t-2}y_{ij,t}$.

Non temporal dependencies which assess the dependencies between paired comparison responses *within* a time point can also be incorporated. These dependencies arise when a particular response to a comparison of two items affects the responses to other paired comparisons. Consider for example the paired comparisons involving the object pairs (O_i, O_j) and (O_i, O_k) , dependency is introduced by the same object O_i involved in both pairs which can be characterized by a further parameter $\theta_{ij,ik|t} := \theta_{i,jk|t}$. Allowing for this type of dependencies the exp-term in (6) has to be augmented by terms like $\theta_{i,jk|t}y_{ij,t}y_{ikt}$. In this case the parameters representing the dependencies can, however, only be interpreted as log-odds ratios in the conditional distribution where conditioning on all previous, intermediate and subsequent Y 's has to be done and this is perhaps not a good way to interpret time series structures (Cox and Wermuth (1994)).

Note that all of the above extensions can easily be incorporated into the model by a straightforward modification of the design matrix \mathbf{X} . Dependencies are represented as two-way interactions and can therefore be constructed by elementwise multiplication of corresponding columns of the response pattern matrix \mathbf{Y} .

We emphasise that the example used is purely illustrative - by converting likert scales to paired comparison form we are discarding information about the absolute rating of each item and retaining only relative comparisons. The resulting analysis therefore displays information on the relative worths of each item rather than the absolute ratings. However, such an analysis is often of great interest. For example, in political surveys which assess the competence of the major party leaders, it is the *relative* worths which is of primary interest, and an analysis of the type proposed here provides a convenient way of summarising

multiple likert scales.

We conclude by conceding that there are still problems to be solved. This model, in common with other methods is only tractable if there are not too many objects to be compared at not too many time points, because the response pattern matrix \mathbf{Y} will quickly become very large for large J and T . The development of efficient numerical techniques for dealing with such problems needs to be a priority in dealing with such data.

References

- Agresti, A. (1992). Analysis of ordinal paired comparison data. *Applied Statistics* 41, 287–297.
- Böckenholt, U. (1988). A Logistic Representation of Multivariate Paired-Comparison Models. *Journal of Mathematical Psychology* 32, 44–63.
- Böckenholt, U. and W. Dillon (1997). Modelling within-subject dependencies in ordinal paired comparison data. *Psychometrika* 62, 411–434.
- Bradley, R. and M. Terry (1952). Rank Analysis of Incomplete Block Designs. I. The Method of Paired Comparisons. *Biometrika* 39, 324–345.
- Buck, N., J. Gershuny, D. Rose, and J. Scott (1994). *Changing Households: The British Household Panel Survey 1990-1992*. Colchester: ESRC Research Centre on Micro-Social Change, University of Essex.
- Cox, D. (1972). The analysis of multivariate binary data. *Applied Statistics* 21, 113–120.
- Cox, D. and N. Wermuth (1994). A note on the quadratic exponential binary distribution. *Biometrika* 81, 403–408.
- Davidson, R. (1970). Extending the Bradley-Terry model to accommodate ties in paired comparison experiments. *JASA* 65, 317–328.
- Davidson, R. and R. Beaver (1977). On extending the Bradley-Terry model to incorporate within-pair order effects . *Biometrics* 33, 693–702.
- Davidson, R. and R. Bradley (1969). Multivariate paired comparisons: The extension of a univariate model and associated estimation and test procedures . *Biometrika* 56, 81–95.
- Dittrich, R., R. Hatzinger, and W. Katzenbeisser (1998). Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. *Applied Statistics* 47, 511–525.
- Dittrich, R., R. Hatzinger, and W. Katzenbeisser (2002). Modelling dependencies in paired comparison experiments . *Computational Statistics and Data Analysis* 40, 39–57.
- Dittrich, R., W. Katzenbeisser, and H. Reisinger (2000). The analysis of rank ordered preference data based on Bradley-Terry Type Models. *OR Spektrum* 22, 117–134.

- Fahrmeier, L. and G. Tutz (1994). Dynamic stochastic models for time-dependent ordered paired comparison systems. *JASA* 89, 1438–1449.
- Fienberg, S. (1979). Log linear representation for paired comparison models with ties and within-pair order effects. *Biometrics* 35, 479–481.
- Francis, B., R. Dittrich, R. Hatzinger, and R. Penn (2002). Analysing ranks using paired comparison methods: an investigation of value orientation in Europe. *Applied Statistics* 51(3), 319–336.
- Francis, B., M. Green, and C. Payne (1993). *The GLIM System. Release 4 Manual*. Oxford: Clarendon.
- Glickman, M. E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Applied Statistics* 48, 377–394.
- Glickman, M. E. (2001). Dynamic paired comparison models with stochastic variances. *Journal of Applied Statistics* 28, 673–689.
- Kousgaard, N. (1976). Models for paired comparisons with ties. *Scand. J. Statist* 3, 1–14.
- Kousgaard, N. (1984). Analysis of a sound field experiment by a model of paired comparisons with explanatory variables. *Scand. J. Statist* 11, 51–57.
- Matthews, J. and K. Morris (1995). An application of Bradley-Terry-type models to the measurement of pain. *Applied Statistics* 44, 243–255.
- Office for National Statistics (1999). *What exactly is the Labour Force Survey?* London: Office for National Statistics.
- Rao, P. and L. Kupper (1967). Ties in paired-comparison experiments: a generalisation of the Bradley-Terry model. *JASA* 62, 194–204.
- Sinclair, C. (1982). GLIM for preference. In R. Gilchrist (Ed.), *Proceedings of the International Conference on Generalised Linear Models*, Volume 14, pp. 164–178. Springer Lecture Notes in Statistics.
- Tutz, G. (1986). Bradley-Terry-Luce models with an ordered response. *Mathematical Psychology* 30, 306–316.