Figure 1

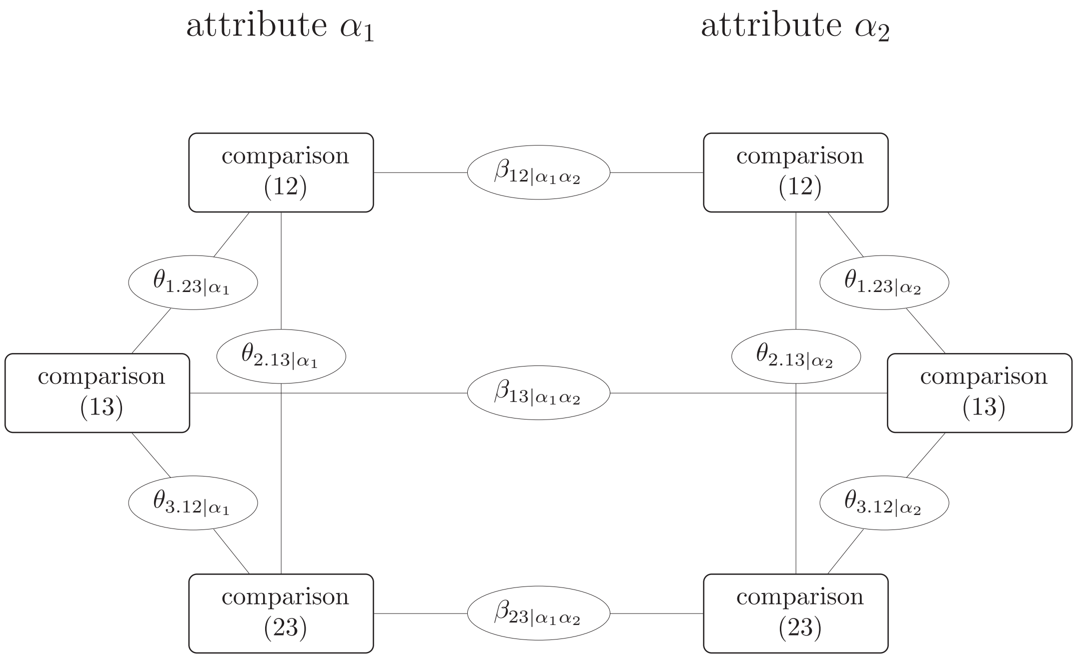

Figure 1: Dependency structure

Figure 2

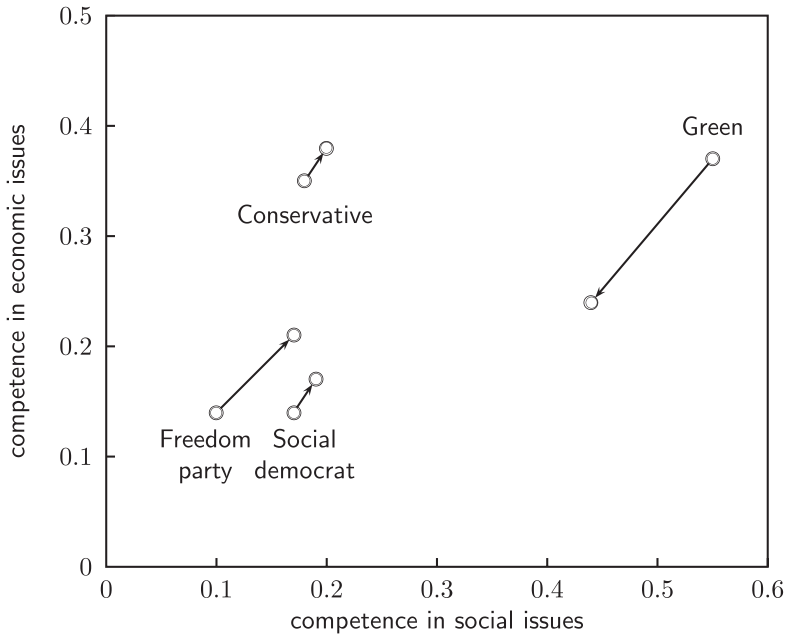

Figure 2: Ascribed competence in social and economic issues

* Manuscript

Modelling Dependency in Multivariate Paired Comparisons: a Log-Linear Approach

Regina Dittrich†

Vienna University of Economics, Austria

Brian Francis

Lancaster University, UK

Reinhold Hatzinger

Walter Katzenbeisser

Vienna University of Economics, Austria

Summary. A log-linear representation of the Bradley-Terry model is presented for multivariate paired comparison data, where judges are asked to compare pairs of objects on more than one attribute. By converting such data to multiple binomial responses, dependencies between the decisions of the judges as well as possible association structures between the attributes can be incorporated in the model, providing an advantage over parallel univariate analyses of individual attributes. The approach outlined gives parameters which can be interpreted as (conditional) log-odds and log-odds ratios. As the model is a generalised linear model, parameter estimation can use standard software, and the GLM framework can be used to test hypotheses on these parameters.

Classification Code: C35

Keywords: Paired comparisons, multivariate Bradley-Terry model, within subject dependencies, multivariate binomial data, log-linear model, generalized linear models, GLIM

1. Introduction

The method of paired comparisons addresses the problem of determining the scale values of a set of J objects $O_1, O_2, \ldots, O_J$ on a preference continuum that is not directly observable. Paired comparisons are judgmental tasks that typically involve repeatedly exposing an individual to a selection of pairs of objects chosen from this set of objects one at a time and asking for a judgment about which element of the pair is preferred.

This sort of experiment results in $\binom{J}{2}$ paired comparisons, say in the pre-defined order

$$(1,2),(1,3),\ldots,(1,J);(2,3),(2,4),\ldots,(2,J);\ldots;(J-1,J)\,, \tag{1}$$

where (i,j) is a shorthand notation for the comparison of objects O_i and O_j. One of the most prominent and well-known models that covers such situations is due to Bradley and Terry (1952). The basic Bradley-Terry (BT-) model is defined by

$$P\{O_i > O_j\} = \frac{\pi_i}{\pi_i + \pi_j}\,,\ P\{O_j > O_i\} = \frac{\pi_j}{\pi_i + \pi_j} \tag{2}$$

†*Address for correspondence:* Regina Dittrich, Vienna University of Economics, Department of Statistics and Mathematics, Augasse 2-6, A-1090 Vienna, Austria. FAX: +43 1 31336 774
E-mail: Regina.Dittrich@wu-wien.ac.at

where $\{O_i > O_j\}$ or $\{O_j > O_i\}$ denote the events that object i or object j is chosen in the comparison of objects i and j. The π's are unknown non-negative parameters, the so called 'worth' parameters, describing the location of the objects on the preference scale which have to be estimated from the observations. To calculate the worth parameters π we have to take into account that the BT-model is invariant under change of scale, and identifiability is usually achieved by the requirement that $\sum_i \pi_i = 1$.

The basic BT-model has been extensively discussed in the literature (for a review cf. e.g. David (1988)) and various extensions have been proposed. To name just a few of those: Ties (Rao and Kupper (1967), Davidson (1970), Kousgaard (1976)); order effects (Davidson and Beaver (1977), Fienberg (1979)); the incorporation of explanatory variables (Kousgaard (1984), Matthews and Morris (1995), Dittrich, Hatzinger and Katzenbeisser (1998), Francis, Dittrich, Hatzinger and Penn (2002)); ordinal paired comparison models (Agresti (1992), Böckenholt and Dillon (1997)).

In all the above work, the objects are compared solely on a single attribute. This paper examines multivariate paired comparisons, where the objects are compared on more than one attribute (Davidson and Bradley (1969), Böckenholt (1988)). For example, a collection of cameras could be compared on picture quality and ease of use. As in the basic BT-model it is assumed that for each attribute α, $\alpha = 1, 2, \ldots, p$, there exists a separate continuum on which the parameters $\pi_{1\alpha}, \pi_{2\alpha}, \ldots, \pi_{J\alpha}$ representing the worth of the objects with regard to the attributes are located. The probability of preferring object i over object j for attribute α is also defined in Bradley-Terry form

$$P\{O_i >_\alpha O_j\} = \frac{\pi_{i\alpha}}{\pi_{i\alpha} + \pi_{j\alpha}}\,, \tag{3}$$

and identifiability is again achieved by setting $\sum_{i=1}^{J} \pi_{i\alpha} = 1$. To fit multivariate BT-models an incomplete contingency table approach due to Imrey, Johnson and Koch (1976) which is based on the Grizzle-Starmer-Koch approach for the analysis of categorical data by linear models (Grizzle, Starmer and Koch (1969)) can be used. Another approach based on a logistic representation (Böckenholt (1988)) can be applied.

In almost all BT-models a more or less explicit assumption is that all decisions of the judges are independent, an assumption which might be questionable at least for the decisions of a given judge: In paired comparison studies, a judge chooses among objects several times, and in such cases, judgements made by the same judge are likely to be dependent. The stochastic nature of the data is now a result of between- and within-subject sources of variation. These possible dependencies should of course be incorporated in the modelling process. Thus the aim of this paper is to present a *log-linear representation* for multivariate paired comparisons, where the main issue is the modelling of various possible dependencies between the decisions of the judges. The statistical modelling of the multivariate paired comparisons will be embedded in the analysis of multiple binomial responses (Cox (1972)). The model presented in this paper can be seen as a generalization of the log-linear model used for modelling dependencies in univariate paired comparisons (Dittrich, Hatzinger, Katzenbeisser (2002)).

In principle two different types of dependencies can be considered: According to the ideas of Böckenholt and Dillon (1997) the association between pairs of paired comparisons can be taken into account. Dependencies between responses are introduced by repeated evaluation of identical objects in a paired comparison experiment. If a judge uses the same standard in comparing objects i and j and objects i and k on attribute α, the assessment of object i is likely to be similar in both comparisons for the given attribute. This similarity in the evaluation of object i might introduce dependencies between the observed responses. Let us call it *between object pairs dependencies*. Consider for example the paired comparisons involving the object pairs (O_i, O_j) and (O_i, O_k) for a given attribute α; dependency is introduced by the same object O_i involved in both pairs which is characterized by a further parameter $\theta_{ij,ik|\alpha} := \theta_{i,jk|\alpha}$. For pairs of paired comparisons that do not involve the same object we set $\theta_{ij,kl|\alpha} = 0$. This case, however, occurs only if $J \geq 4$. Therefore two pairwise responses are regarded as independent when they are based on two nonoverlapping sets of object pairs. A second type of dependencies can be defined for a given comparison over (two) attributes, so called *within attribute dependencies*. For example there could be an association between the outcome of the comparison of object i with object j with respect to the two attributes α_1 and α_2. This possible association will be represented by the parameter $\beta_{ij|\alpha_1\alpha_2}$.

The proposed log-linear approach has several advantages: (i) First, modelling is done within the Generalised Linear Model (GLM-) framework, thus parameter estimates can be obtained by using standard software, e.g. GLIM (Francis et al. (1993)). Moreover various hypotheses about the parameters of the model can easily be tested within the GLM framework by comparing deviance differences of the involved (nested) models. (ii) The second advantage of the log-linear approach is that both types of dependencies can be incorporated into the analysis in the usual GLM way as two-way interaction. Moreover, this specification allows in principle also that higher order dependencies, i.e. dependencies involving more than two objects, can also be taken into account. Therefore, this simultaneous modelling gives an advantage over parallel univariate analysis of single attributes. (iii) The parameters of interest can as usual in the GLM framework be interpreted in terms of log-odds and log-odds ratios, however in a conditional sense.

2. A log-linear approach for multivariate paired comparisons

In this section we will present a log-linear formulation of the multivariate paired comparison experiment. With regard to the aim of formulating a log-linear representation a multiplicative specification, rather than an additive specification (Bahadur (1961)), of the underlying joint probability distribution is used.

2.1. Parameter estimation

Modelling starts with the following representation of the multivariate paired comparison experiment which has the advantage that this approach can easily be adapted to incorporate both types of dependencies.

Consider N judges who independently undergo a multivariate paired comparison experiment, where it is assumed that each judge compares each pair of objects on all p attributes. The result for each comparison can be represented by random variables $Y_{ij\alpha}$, $i, j = 1, 2, \ldots, J, i < j$, $\alpha = 1, 2, \ldots, p$ where

$$Y_{ij\alpha} = \begin{cases} 1 & \text{if } O_i > O_j \text{ on attribute } \alpha\,, \\ -1 & \text{if } O_j > O_i \text{ on attribute } \alpha\,. \end{cases} \tag{4}$$

Hence, for a given judge the experiment results in one of $2^{p\binom{J}{2}} = \ell$ possible response pattern vectors $\mathbf{y}_i$, $i = 1, 2, \ldots, \ell$, a vector of $p\binom{J}{2}$ elements consisting entirely of $\{1, -1\}$, also in the pre-defined order (1). In general the response pattern vector can be written as

$$\mathbf{y} = (y_{121}, y_{122}, \ldots, y_{12p}\,; y_{131}, y_{132}, \ldots, y_{13p}\,; \ldots;\, y_{J-1,J,1}, y_{J-1,J,2}, \ldots, y_{J-1,J,p})\,,$$

with $y_{ij\alpha} \in \{1, -1\}$ according to (4). The order in which the responses are observed could be the standard order of a $2^{p\binom{J}{2}}$ factorial main effects only design; this means that the rightmost y varying the fastest and the leftmost y varying the slowest with all possible combinations being produced. A few response pattern vectors are then given by

$$\begin{aligned}
\mathbf{y}_1 &= (1, 1, \ldots, \quad 1, \quad 1),\\
\mathbf{y}_2 &= (1, 1, \ldots, \quad 1, -1),\\
\mathbf{y}_3 &= (1, 1, \ldots, -1, \quad 1),\\
\mathbf{y}_4 &= (1, 1, \ldots, -1, -1),\\
&\vdots\\
\mathbf{y}_\ell &= (-1, -1, \ldots, -1, -1).
\end{aligned}$$

In order to get a log-linear representation for the multivariate paired comparison experiment that can be extended to account for the mentioned dependencies we have to model the joint distribution of the the random variable

$$\mathbf{Y} = \{Y_{121}, Y_{122}, \ldots, Y_{12p}; Y_{131}, Y_{132}, \ldots, Y_{13p}; \ldots; Y_{J-1,J,1}, Y_{J-1,J,2}, \ldots, Y_{J-1,J,p}\}\,.$$

Example: To illustrate the log-linear approach consider the special case $J = 3$ and $p = 2$. According to the multiplicative specification for multivariate binary data due to Cox (1972), we specify the joint probability distribution for the random variable $\mathbf{Y}$ in the following way: by using Sinclair's reparameterization (Sinclair(1982)) of the Bradley-Terry specification (3) for a single comparison

$$P\{Y_{ij\alpha} = y_{ij\alpha}\} = \frac{1}{\sqrt{\pi_{i\alpha}/\pi_{j\alpha}} + \sqrt{\pi_{j\alpha}/\pi_{i\alpha}}} \left(\frac{\sqrt{\pi_{i\alpha}}}{\sqrt{\pi_{j\alpha}}}\right)^{y_{ij\alpha}}, \; y_{ij\alpha} \in \{-1, 1\} \tag{5}$$

we obtain the joint distribution analogously to Dittrich, Hatzinger and Katzenbeisser (2002):

$$\begin{aligned}
P\{Y_{121} &= y_{121}, Y_{122} = y_{122}; Y_{131} = y_{131}, Y_{132} = y_{132}; Y_{231} = y_{231}, Y_{232} = y_{232}\} =\\
&= \Delta \left(\frac{\sqrt{\pi_{11}}}{\sqrt{\pi_{21}}}\right)^{y_{121}} \left(\frac{\sqrt{\pi_{12}}}{\sqrt{\pi_{22}}}\right)^{y_{122}} \left(\frac{\sqrt{\pi_{11}}}{\sqrt{\pi_{31}}}\right)^{y_{131}} \left(\frac{\sqrt{\pi_{12}}}{\sqrt{\pi_{32}}}\right)^{y_{132}} \left(\frac{\sqrt{\pi_{21}}}{\sqrt{\pi_{31}}}\right)^{y_{231}} \left(\frac{\sqrt{\pi_{22}}}{\sqrt{\pi_{32}}}\right)^{y_{232}} \times\\
&\quad \exp\{\theta_{1,23|1}\, y_{121} y_{131}\} \exp\{\theta_{1,23|2}\, y_{122} y_{132}\} \times\\
&\quad \exp\{\theta_{2,13|1}\, y_{121} y_{231}\} \exp\{\theta_{2,13|2}\, y_{122} y_{232}\} \times\\
&\quad \exp\{\theta_{3,12|1}\, y_{131} y_{231}\} \exp\{\theta_{3,12|2}\, y_{132} y_{232}\} \times\\
&\quad \exp\{\beta_{12|12}\, y_{121} y_{122}\} \exp\{\beta_{13|12}\, y_{131} y_{132}\} \exp\{\beta_{23|12}\, y_{231} y_{232}\}\,,
\end{aligned} \tag{6}$$

where Δ is a normalizing constant in order to make the probabilities in (6) sum up to unity. Figure 1 illustrates the various types of dependencies considered in this example.

FIGURE 1 ABOUT HERE
Figure 1: Dependency structure

Furthermore, let N_i be the random variable

$$N_i = \text{number of times where response pattern vector } \mathbf{y}_i,\ i = 1, 2, \ldots, \ell,\ \text{occurs},$$

then the N_i's are multinomially distributed with $N = \sum_i N_i$, and probabilities given in (6). Now let m_i be the expectation of N_i, i.e. the expected number of times where response pattern vector $\mathbf{y}_i$ occurs. Hence we base our estimation procedure upon simple multinomial sampling. Thus we obtain for the logarithm of the expectation m_i of the random variables N_i the following linear representation

$$\begin{aligned} \ln m \;=\; & \gamma + \lambda_{11}(y_{121} + y_{131}) + \lambda_{12}(y_{122} + y_{132}) + \\ & \lambda_{21}(y_{231} - y_{121}) + \lambda_{22}(y_{232} - y_{122}) + \\ & \lambda_{31}(-y_{131} - y_{231}) + \lambda_{32}(-y_{132} - y_{232}) + \\ & \theta_{1,23|1}\, y_{121}y_{131} + \theta_{1,23|2}\, y_{122}y_{132} + \\ & \theta_{2,13|1}\, y_{121}y_{231} + \theta_{2,13|2}\, y_{122}y_{232} + \\ & \theta_{3,12|1}\, y_{131}y_{231} + \theta_{3,12|2}\, y_{132}y_{232} + \\ & \beta_{12|12}\, y_{121}y_{122} + \beta_{13|12}\, y_{131}y_{132} + \beta_{23|12}\, y_{231}y_{232}\,, \end{aligned} \tag{7}$$

where $\gamma = \ln \Delta$ and $\lambda_{i\alpha} = \frac{1}{2} \ln \pi_{i\alpha}$. For example, the log-linear representation for the expectation of N_1, the expected number of times, where the response pattern vector $\mathbf{y}_1 = (1, 1; 1, 1; 1, 1)$ occurs is given by

$$\begin{aligned} \ln m_1 \;=\; & \gamma + 2\lambda_{11} + 2\lambda_{12} - 2\lambda_{31} - 2\lambda_{32} + \\ & \theta_{1,23|1} + \theta_{1,23|2} + \theta_{2,13|1} + \theta_{2,13|2} + \theta_{3,12|1} + \theta_{3,12|2} + \\ & \beta_{12|12} + \beta_{13|12} + \beta_{23|12}\,. \end{aligned}$$

Model (7) is a Generalised Linear Model and the parameters can easily be estimated by standard software using, e.g. by GLIM using Poisson error and a log-link. The design matrix for the case under considerations, is given by

$$\mathbf{X} = (\mathbf{1}, \mathbf{YA}, \mathbf{W})\,, \tag{8}$$

where $\mathbf{1}$ is a column vector consisting of 1. The $(2^6 \times 6)$-matrix $\mathbf{Y}$, the response pattern matrix, is the design matrix for a 2^6 main effects only design in standard order (in fact the rows of $\mathbf{Y}$ are the response pattern vectors $\mathbf{y}_i$ and is given by

$$\mathbf{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \\ \vdots \\ \mathbf{y}_\ell \end{pmatrix},$$

and

$$\mathbf{A} = \mathbf{B} \otimes \mathbf{I}_2\,,$$

where $\mathbf{B}$ is the (3×3) paired comparison design matrix (Böckenholt and Dillon (1997)). Each column of this matrix corresponds to one of the 3 objects, and each row to one of the $\binom{3}{2} = 3$ paired comparisons (in the order given in (1))

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} ;$$

$\mathbf{I}_2$ denotes the (2×2) identity matrix, and the index refers to the number of attributes, i.e, $p = 2$. The matrix $\mathbf{W}$ is given by $\mathbf{W} = (\mathbf{W}_1, \mathbf{W}_2)$, where the columns of $\mathbf{W}_1$ represent the between object pairs dependencies, and the columns of $\mathbf{W}_2$ represents the within attribute dependencies, according to our previous definition. Therefore, all columns of the matrix $\mathbf{W}$ can be interpreted as representing two-way interactions between responses corresponding to columns of $\mathbf{Y}$ and can easily be constructed in the usual GLM way as two-way interactions by elementwise multiplication of suitable columns of the response pattern matrix $\mathbf{Y}$. Let $\mathbf{y}_{ij\alpha}$ be the columns of the response pattern matrix, corresponding to the comparison of objects i and j on attribute α, i.e. $\mathbf{Y} = (\mathbf{y}_{121}, \mathbf{y}_{122}; \mathbf{y}_{131}, \mathbf{y}_{132}; \mathbf{y}_{231}, \mathbf{y}_{232})$, the matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ are given by

$$\mathbf{W}_1 = (\mathbf{y}_{121} \odot \mathbf{y}_{131}, \mathbf{y}_{122} \odot \mathbf{y}_{132}, \mathbf{y}_{121} \odot \mathbf{y}_{231}, \mathbf{y}_{122} \odot \mathbf{y}_{232}, \mathbf{y}_{131} \odot \mathbf{y}_{231}, \mathbf{y}_{132} \odot \mathbf{y}_{232}) ,$$

and

$$\mathbf{W}_2 = (\mathbf{y}_{121} \odot \mathbf{y}_{122}, \mathbf{y}_{131} \odot \mathbf{y}_{132}, \mathbf{y}_{231} \odot \mathbf{y}_{232}) ,$$

where $\odot$ represents the elementwise (Hadamard) product of the corresponding columns.

The design matrix for the simplest case as shown in the example can be generalized for more than three objects and more than two attributes in an obvious way. For the general case with J objects and p attributes the design matrix is analogously to (8) given by

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

where the matrix $\mathbf{Y}$ is the $(2^{p\binom{J}{2}} \times p\binom{J}{2})$ design matrix for a $2^{p\binom{J}{2}}$ main effects only design in standard order, $\mathbf{A} = \mathbf{B} \otimes \mathbf{I}_p$, where $\mathbf{B}$ is the $(\binom{J}{2} \times J)$ paired comparison design matrix given by

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

and $\mathbf{I}_p$ is the corresponding identity matrix of order p, i.e. the number of attributes. Care has only to be taken when specifying the matrix $\mathbf{W}$ because not all possible interactions between columns of $\mathbf{Y}$ are meaningful according to the definition of the origin causing the dependencies as there are nonoverlapping sets of object indices, when $J \geq 4$. The log-linear model is overparameterized and one has to impose some restrictions on the parameters. In this paper we use the standard GLIM-restrictions, and aliase the λ parameters accordingly.

2.2. *Parameter interpretation*

Even in this small example of three objects and two attributes, there are already six association parameters. An interesting question might be whether all association parameters are needed in the specification of the model. For example one could be interested in the hypothesis $\theta_{1,23|\alpha} = \theta_{1,24|\alpha} = \cdots = \theta_{1|\alpha}$. Another, more restrictive hypothesis, would be $\theta_{1|1} = \theta_{1|2} = \cdots = \theta_{1|p} = \theta_{1|\cdot}$. Furthermore, another hypothesis is $\theta_{1,23|1} = \theta_{1,23|2} = \cdots = \theta_{1,23|p}$. The most restrictive hypothesis would be that all association parameters are zero, in which case the independence model is achieved. It might also be of interest whether the β-parameters can be restricted in some interesting way. An obvious advantage of this GLM-approach is that all mentioned hypotheses can easily be tested in the usual GLM way by comparing deviances of the associated nested models.

The interpretation of the parameters of model (7) is best seen by considering conditional distributions, because as Cox pointed out, conditional distributions from this class of models have a simple form, marginal probabilities do not. Consider for example the log-odds in favour of Y_{121} conditional on all other $Y_{ij\alpha}$:

$$\ln \frac{P\{Y_{121} = 1|\mathbf{Y}^-\}}{P\{Y_{121} = -1|\mathbf{Y}^-\}} = 2(\lambda_{11} - \lambda_{21}) + 2(\theta_{1,23|1} y_{131} + \theta_{2,13|1} y_{231}) + 2\beta_{12|12} y_{122}\,,$$

where $\mathbf{Y}^-$ denotes the random vector $\mathbf{Y}$ without the element Y_{121}. Therefore, for given attribute $\alpha = 1$, the log-odds in favour of object O_1 are not only determined by the parameters of the involved objects, as in the basic BT-model, but additionally (i) all θ's have to be taken into account which represents interactions between those pairs of paired comparisons which involve the pair (O_1, O_2), i.e. the pairs $((O_1, O_2), (O_1, O_3))$ and $((O_1, O_2), (O_2, O_3))$, and (ii) there is perhaps also an effect of the within attribute dependency. Therefore, for two evenly matched objects, and assuming no within attribute dependency, i.e. $\beta_{12|12} = 0$, there is an advantage in preferring object O_1 over object O_2 regarding attribute $\alpha = 1$ if $\theta_{1,23|1} y_{131} + \theta_{2,13|1} y_{231} > 0$, which for example is given when $\theta_{1,23|1}$ and $\theta_{2,13|1}$ are positive and $y_{131} = y_{231} = 1$. Note that this advantage is caused solely by the assumed between object pairs dependency.

The θ-parameters are proportional to log-odds ratios describing the association between two Y's in the conditional distribution of two paired comparisons, i.e. two Y's, given the others. For example, the parameter $\theta_{1,23|1}$ is proportional to the log-odds ratio in the conditional distribution of $\{Y_{121}, Y_{131}\}$ given $\mathbf{Y}^-$, where $\mathbf{Y}^-$ now denotes the random vector $\mathbf{Y}$ without the elements $\{Y_{121}, Y_{131}\}$:

$$\ln \frac{P\{Y_{121} = 1, Y_{131} = 1|\mathbf{Y}^-\} P\{Y_{121} = -1, Y_{131} = -1|\mathbf{Y}^-\}}{P\{Y_{121} = 1, Y_{131} = -1|\mathbf{Y}^-\} P\{Y_{121} = -1, Y_{131} = 1|\mathbf{Y}^-\}} = 4\,\theta_{1,23|1}\,.$$

A positive association parameter $\theta_{1,23|1}$ indicates that the judges will rather be consistent (positive or negative) within their decisions between objects O_1 and O_2 for a given attribute. Following Böckenholt and Dillon (1997) the θ parameters can be interpreted as an indicator of a stimulus identity effect that reflects the degree of similarity or consistency in the two assessments of the common object with regard to the attribute. In general, however, the interpretation of the sign of the association parameters θ is not so straight forward, because it depends on the object indices involved in the pairs of paired comparisons. If the common index i is the smallest or the largest of the involved indices i, j, k than a positive parameter suggests consistency of the decisions, but in all other cases a negative parameter indicates consistency. This is caused by the ordering of the two-dimensional tables representing the two-dimensional conditional distributions of the random variables $Y_{ij|\alpha}$ and $Y_{ik|\alpha}$, as will be shown in the following example.

A similar interpretation can be given for the β-parameters. For example $\beta_{12|12}$ is proportional to the log-odds ratio in the conditional distribution of $\{Y_{121}, Y_{122}\}$ given all other

Y's:

$$\ln \frac{P\{Y_{121}=1, Y_{122}=1|\mathbf{Y}^-\}P\{Y_{121}=-1, Y_{122}=-1|\mathbf{Y}^-\}}{P\{Y_{121}=1, Y_{122}=-1|\mathbf{Y}^-\}P\{Y_{1211}=-1, Y_{122}=1|\mathbf{Y}^-\}} = 4\,\beta_{12|12}\,.$$

A positive (negative) parameter $\beta_{12|12}$ indicates therefore a positive (negative) association between both attributes within the comparison of objects 1 and 2, given all other decisions.

3. An Application

In this section we will illustrate our approach by means of an example. A paired comparison experiment was carried out at the Vienna University of Economics during the summer term 2001, after the election in the year 1999. 266 mainly first-year students were asked to state a preference about the leaders of the four political parties represented in the Austrian Parliament. In this experiment two attributes, namely *competence in social issues, s,* and *competence in economic issues, e,* were taken into account. The leaders of the four political parties are denoted by the abbreviations S, the head of the social democrats, F, the leader of the so-called freedom party , C, the leader of the conservatives, and G the head of the green (ecologic) party (sorted by the number of votes reached in the election 1999). Two questions for each pairwise comparison were asked, namely which of two party leaders has more competence in social issues and more competence in economic issues, respectively.

As a first step the independence model for four items (=party leaders) was fitted by GLIM and the results are given in Table 3 in the appendix. From the estimates of the initial λ parameters the worth parameters $\hat{\pi}$ for the four party leaders with respect to the two attributes can be calculated by

$$\hat{\pi}_{i\alpha} = \frac{\exp\{2\hat{\lambda}_{i\alpha}\}}{\sum_{\ell}\exp\{2\hat{\lambda}_{\ell\alpha}\}}$$

and are shown in the following table:

	$\hat{\pi}_s$	$\hat{\pi}_e$
G	0.55	0.37
C	0.18	0.35
S	0.17	0.14
F	0.10	0.14

From the independence model one would conclude that G, the leader of the Green party, is thought to have the most competence in social and economic issues. However, the distance to C in the economic attribute is not significant; an approximate 0.95%-confidence interval for the initial parameter λ_{G_e} is given by $[-0.072, 0.109]$. S and F are quite similar, although S is seen to have a higher social competence, the corresponding approximate 0.95%-confidence interval for λ_{S_s} is given by $[-0.106, 0.074]$. One has to be careful in interpreting the overall fit of the independence model because there are numerous random zeros in the underlying contingency table, which inflates the residual degrees of freedom. Therefore we consider deviance changes of hierarchical models only.

The next step is to incorporate dependencies between the decisions of the students into the model. Therefore we fitted the 'full' model allowing for *between objects pair dependencies* as well as *within attribute dependencies* and the results are shown in Table 4 in the appendix. Comparing the independence model with the 'full' model, one can see a tremendous improvement of the fit, with a change of deviance 813.3 on 30 degrees of freedom. This suggests that the independence model is not appropriate for this data set. It is interesting that the worth parameters for the party leaders based on this dependence model change:

	$\hat{\pi}_s$	$\hat{\pi}_e$
G	0.44	0.24
C	0.20	0.38
S	0.19	0.17
F	0.17	0.21

The changes in the model parameters are displayed in Figure 2, where the arrows represent the direction of the changes, with starting points given by the worth parameters calculated on the basis of the independence model.

FIGURE 2 ABOUT HERE

Figure 2: Ascribed competence in social and economic issues

The effect of the incorporation of the dependencies leads to, at least in this example, more similarity between the party leader, which can be seen in Figure 2. The ascribed competence of G, the head of the green party, changes most. Compared to the initial model he is seen to be less competent in both attributes, but still he is ascribed the highest competence in social issues of all party leaders; an approximate 0.95% confidence interval for the corresponding initial social competence parameter is given by [0.275; 0.519]. C gains in both attributes and overtakes G in the ascribed competence in economic issues.

Concerning the interaction parameters θ they have positive as well as negative signs. Nevertheless, all those interaction parameters indicate that there is a tendency towards consistent decisions. Consider for example the parameters $\theta_{S,GF|s}$ and $\theta_{F,SC|s}$. To make it easier to read we will denote the choices e.g. $\{Y_{SFs} = 1\}$, if the first leader (S) is chosen, by $SF >_s S$ and and the choice $\{Y_{SFs} = -1\}$, if the second leader (F) is chosen, by $SF >_s F$.

The (estimated) parameter $\hat{\theta}_{S,GF|s}$ implies that the (estimated) choice probabilities can be written as:

$$P\{SG >_s S, SF >_s S|Y^-\}\ P\{SG >_s G, SF >_s F|Y^-\} =$$
$$\exp\{4 \cdot 0.3233\} \quad P\{SG >_s S, SF >_s F|Y^-\}\ P\{SG >_s G, SF >_s S|Y^-\}\,.$$

The responses on the left hand side are consistent decisions about S, because $SG >_s S$ and

$SF >_s S$ both mean that S, the first party leader in the comparisons is seen to have more competence in social issues (s) when compared to F and G while $SF >_s F$ and $SG >_s G$ mean that in both cases S is seen to have less competence in social issues when compared to F and G. Whereas the responses in the other side are inconsistent decisions about S, because S is seen to have more competence in social issues than F ($SF >_s S$) but less competence in social issues than G ($SG >_s G$). This illustrates that in this case the chance for a consistent decision ($\{SF >_s S, SG >_s S\}\{SF >_s F, SG >_s G\}$) of the students for either always choosing S to be the party leader with more competence in social issues or never choosing S when compared to the party leaders F or G is $\exp\{1.2932\} = 3.64$ times higher than the probability for choosing S just once when compared to the party leader F or G. In other words one is more likely either to prefer the social democrat to both the conservative and freedom party leader, or to prefer the freedom and conservative party leader to the social democrat.

On the other hand for $\theta_{F,SC|s}$ we get

$$P\{SF >_s S\,, FC >_s F|Y^-\}\ P\{SF >_s F, FC >_s C|Y^-\} =$$
$$\exp\{-4 \cdot 0.6853\}\ P\{SF >_s S, FC >_s C|Y^-\}\ P\{SF >_s F, FC >_s F|Y^-\}\,.$$

In this case consistency about the leader F is indicated by $\{SF >_s S, FC >_s C\}\{SF >_s F, FC >_s F\}$ which is now on the right hand side of the formula above, and therefore the chance for consistency is $\frac{1}{\exp\{-4\cdot 0.6853\}} \approx 16$ times higher than the chance for inconsistency.

In order to obtain a more parsimonious model, we tried to simplify the θ's, according to the suggestions made at the end of Section 2. However, in general this is not possible in this example. Consider for example the hypothesis $\theta_{S,GF|s} = \theta_{S,GC|s} = \theta_{S,FC|s} = \theta_{S|s}$. This hypothesis has to be rejected, change in deviance is 18.7 on 2 df. This is the reason why we did not simplify the model further with respect to the θ parameters.

It is also remarkable that all β-parameters are positive and approximately of the same size. Therefore the model can further be simplified by substituting six parameters by only one β, with an estimated value of 0.3385 (s.e.=0.0297), change in deviance is 2.03 on 5 df. This can be interpreted as a positive association between the competence in social- and economic issues.

Summarizing the results of this experiment one could conclude that it is important to consider interaction parameters representing dependencies between the decisions of the judges on the one hand and attributes on the other hand, otherwise one might get biased estimates for the worth parameters.

4. Acknowledgements

Parts of this work was supported by the UK ESRC under the National Centre for Research Methods initiative (RES-576-25-5020 - Developing Statistical Modelling in the Social Sciences).

5. References

Agresti, A., 1990. Categorical Data Analysis. J.Wiley, New York.
Agresti, A., 1992. Analysis of ordinal paired comparison data. Applied Statistics, 41, 287-297.
Bahadur, R. R., 1961. A representation of the joint distribution of responses to n dichotomous items. In: Solomon, H. (Eds.), Studies in Item Analysis and Prediction. Stanford University Press, Stanford, pp. 158-176,
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. A. and M. E. Terry, 1952. Rank analysis of incomplete block designs I. The method of paired comparisons. Biometrika, 39, 324-345.
Cox, D., 1972. The analysis of multivariate binary data. Applied Statistics, 21, 113-120.
David, H. A., 1988. The Method of Paired Comparisons, 2nd ed. Griffin, London.
Davidson, R. R., 1970. Extending the Bradley-Terry model to accommodate ties in paired comparison experiments. Journal of the American Statistical Association, 65, 317-328.
Davidson, R. R. and R. A. Bradley, 1969. Multivariate paired comparisons: The extensions of a univariate model and associated estimation and test procedures. Biometrika, 56, 81-95.
Davidson, R. R. and R. J. Beaver, 1977. On extending the Bradley-Terry model to incorporate within-pair order effects. Biometrics, 33, 693-702.
Dittrich R., R. Hatzinger and W. Katzenbeisser, 1998. Modelling the effect of subject-specific covariates in paired comparison studies with an application to university ranking. Applied Statistics, 47, 511-525.
Dittrich, R., R. Hatzinger, and W. Katzenbeisser, 2002. Modelling dependencies in paired comparison experiments. A log-linear approach. Computational Statistics & Data Analysis, 40, 39-57.
Fienberg, S. E., 1979. Log linear representation for paired comparison models with ties and within-pair order effects. Biometrics, 35, 479-481.
Francis, B. et.al (Eds.) 1993. The GLIM System. Release 4 Manual. Clarendon Press, Oxford.
Francis B. R. Dittrich , R. Hatzinger and R. Penn, 2002. Analysing partial ranks by using smoothed paired comparison methods: an investigation of value orientation in Europe. Applied Statistics, 51, 319-336.
Grizzle, J. E., Starmer, C. F. and G. G. Koch, 1969. Analysis of categorical data by linear models. Biometrics, 25, 489-504.
Imrey, P. B., W. D. Johnson, and G. G. Koch, 1976. An incomplete contingency table approach to paired-comparison experiments. Journal of the American Statistical Association, 71, 614-623.
Kousgaard, N., 1976. Models for paired comparisons with ties. Scandinavian Journal of Statistics, 3, 1-11.
Kousgaard, N., 1984. Analysis of a sound field experiment by a model of paired comparisons with explanatory variables. Scandinavian Journal of Statistics, 11, 243-255.
Matthews, J. N. S. and K. P. Morris, 1995. An application of Bradley-Terry-type models to the measurement of pain. Applied Statistics, **44**, 243-255.
Rao, P. V. and L. L. Kupper, 1967. Ties in paired-comparison experiments: a generalisation of the Bradley-Terry model. Journal of the American Statistical Association, 62, 194-204.

Sinclair, C. D., 1982. GLIM for preference. In: Gilchrist, R. (Eds.): GLIM 82. Proceedings of the International Conference on Generalised Linear Models. Springer Lecture Notes in Statistics, 14. pp. 164-178.

6. Appendix

Table 1. Estimates of the λ parameters of the independence multivariate model

```
scaled deviance =   1444.3 at cycle 6
    residual df =   4089
```

	estimate	s.e.	parameter
1	-0.0165	0.04512	S_s
2	-0.4517	0.04643	S_e
3	0.5760	0.04997	G_s
4	0.0186	0.04525	G_e
5	-0.2771	0.04654	F_s
6	-0.4414	0.04633	F_e
7	0.0000	aliased	C_s
8	0.0000	aliased	C_e

Table 2. Parameter estimates for the multivariate dependence model

```
scaled deviance = 630.98 (change = -813.3) at cycle 8
    residual df = 4059 (change = -30 )
```

	estimate	s.e.	parameter
1	-0.0316	0.0616	S_s
2	-0.3889	0.0552	S_e
3	0.3970	0.0610	G_s
4	-0.2252	0.0534	G_e
5	-0.0758	0.0615	F_s
6	-0.3025	0.0565	F_e
7	0.0000	aliased	C_s
8	0.0000	aliased	C_e
9	0.3233	0.1244	THETA S, GF|s
10	0.2024	0.1175	THETA S, GC|s
11	0.9171	0.1506	THETA S, FC|s
12	-0.3229	0.1288	THETA G, SF|s
13	-0.3842	0.1194	THETA G, SC|s
14	0.5094	0.1171	THETA G, FC|s
15	0.5350	0.1221	THETA F, SG|s
16	-0.6853	0.1540	THETA F, SC|s
17	-0.4377	0.1165	THETA F, GC|s
18	0.6618	0.1123	THETA C, SG|s
19	0.6389	0.1561	THETA C, SF|s
20	0.1578	0.1165	THETA C, GF|s
21	0.3106	0.0986	THETA S, GF|e
22	0.2950	0.0973	THETA S, GC|e
23	0.5900	0.1157	THETA S, FC|e
24	-0.3354	0.1035	THETA G, SF|e
25	-0.3425	0.0993	THETA G, SC|e
26	0.7671	0.1243	THETA G, FC|e
27	0.4372	0.0966	THETA F, SG|e
28	-0.4255	0.1170	THETA F, SC|e
29	-0.4611	0.1280	THETA F, GC|e
30	0.3615	0.0959	THETA C, SG|e
31	0.5573	0.1172	THETA C, SF|e
32	0.4304	0.1272	THETA C, GF|e
33	0.3997	0.0803	BETA SG|se
34	0.3526	0.0738	BETA SF|se
35	0.3328	0.0810	BETA SC|se
36	0.3802	0.0880	BETA GF|se
37	0.3216	0.0831	BETA GC|se
38	0.2584	0.0726	BETA FC|se