MODELLING PARALLEL STREAMS OF OBSERVATIONS WITH MULTINOMIAL RESPONSE MODELS

1990

Pieter M. Kroonenberg*, Department of Education, Leiden University
Albert Verbeek, Department of Sociology, University of Utrecht

Abstract

The behaviours of a child and a therapist are scored every five seconds, which gives two parallel streams of observations. These are modelled with Markov chains, which can be recast as multinomial response models, which can be analysed with standard software.

Two substantive problems receive special attention. Firstly, we will examine autodependence and crossdependence. Secondly, at five-second intervals, most of the behaviour is constant between two succesive observations. Hence there is a special interest in modelling changes in behaviour, that is off-diagonal transition probabilities. This leads to multinomial response models for incomplete tables. With standard packages such analyses can be done, but there are errors, traps, and pitfalls along the way. For this reason we compare BMDP4F, SAS CATMOD, and SPSS LOGLINEAR.

1. Introduction

In the behavioural sciences a commonly occurring data type is that of parallel streams of categorically coded behaviour of two partners, sometimes called sequential dyadic data. In this paper we will model such streams, in particular the combined scores of several child-therapist pairs, whose behaviours were scored every five seconds. Two substantive problems receive special attention. Firstly, we will evaluate autodependence (behaviour predicted by own previous behaviour) and crossdependence (behaviour predicted by previous behaviour of the partner). This can be done in a fairly routine way with multinomial response models using standard software packages. Secondly, at five-second intervals, most of the behaviour is constant between two succesive observations. Hence special care has to be taken to model changes in behaviour, which leads to multinomial response models applied to incomplete contingency tables. This is much less routine, as will be shown below.

Especially in psychology parallel streams of behaviour have been studied intensively, and some of the relevant papers are Allison and Liker (1982), Budescu (1984), Wasserman and Iacobucci (1987), Iacobucci and Wasserman (1988), and their references. Most of the papers in the field, however, only deal with binomial response models, whereas our dependent variables are four category ones. In the present paper a balance is sought between application and methodology, but a full methodological paper is currently in preparation.

2. The observational design

Behaviour of both child and therapist have been scored every five seconds during approximately one-hour sessions. Over 70,000 observations were made on 117 pairs (81 children and 4 therapists). Even though important for some substantive questions, we will pay no attention to differences between the various child-therapist combinations, but treat the data as if they came from one single child-therapist combination. The behaviours of the child will be denoted by C, and that of the therapist by T. It will be convenient to denote current behaviour by $C_0 = C$, and $T_0 = T$, behaviour one interval ago by $C_{-1} = \text{lag}(C)$, and $T_{-1} = \text{lag}(T)$, etc. From C and T and their lagged variables we have to build contingency tables for the analysis. Suppose we want to build the table $C\text{x}C_{-1}\text{x}T_{-1}\text{x}C_{-2}\text{x}T_{-2}$. Then we have to create the variables lag(C), lag(lag(C)), lag (T), and lag(lag(T)). One could do this for each different pair child-therapist separately, but we will stack all time series, separating two series by a pseudo-observation with missing values. Hence an observation from the beginning of a time series that contains a lagged variable belonging to another time series will always contain a lagged variable with a missing value. Such cases will be automatically eliminated when we build the contingency table. Thus we have created a (71,276 + 104 - 1) by 5 matrix, from which we can build the frequency table $C\text{x}C_{-1}\text{x}T_{-1}\text{x}C_{-2}\text{x}T_{-2}$, ignoring cases for which any variable is missing.

The original scores were coded using about 34 categories, which were recoded to four categories for each of the partners: NonPlay, Play Preparation, Functional Play, and Imagery Play; for details see Harinck and Hellendoorn, 1987; Hellendoorn, Kroonenberg, and Harinck (1990). Thus C and T and their lagged variables are categorical variables with four categories.

3. The model

We want to build a model that predicts the present behaviours C and T from past behaviour. More precisely, we want to model the probabilities $\Pr(C = i)$ and $\Pr(T = j)$ for $i, j = 1..4$ conditionally upon C_{-1}, T_{-1}, C_{-2}, T_{-2}, etc.. We will refer to p, the vector of values of these variables (denoted with lower case letters), as a **condition** or **pattern**. Suppose we want to condition on C_{-1},

---

*Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. We thank Frits Harinck and Joop Hellendoorn for the permission to use their data set as an illustration and an inspiration.

In J.R. Mathieu (Ed.). Proceedings of the Fifth International Workshop on Statistical Modeling (pp. 116 - 123). Toulouse, France: Laboratoire de Statistique et Probabilités, Université Paul Sabatier.

T_{-1}, C_{-2}, and T_{-2}, then the condition is the pattern $p = (c_{-1}, t_{-1}, c_{-2}, t_{-2})$. Thus the probabilities we want to model are

$$\pi_{C,pi} =_{\text{def}} \Pr(C = i \mid \text{ given condition } p) \text{ for } i = 1..4 \tag{1a}$$

and,

$$\pi_{T,pj} =_{\text{def}} \Pr(T = j \mid \text{ given condition } p) \text{ for } j = 1..4 \tag{1b}$$

Such probabilities are also called 'transition probabilities'; they give the probability of the transition to state i (or j) at the next time point, given the condition p. Restricting ourselves to (1a), we assume that $\pi_{C,pi}$ can be predicted from the previous state and some past states. This is precisely the Markov stationarity assumption, in other words our model is a Markov chain model. On how many variables we have to condition, i.e. the order of the Markov chain, will be evaluated by means of the data. Within the limits allowed by the substantive theory, we are looking for a reasonably well fitting model, which is as simple as possible.

3.1. The parameter model

As predictors in this model we may use the same variables that were used for conditioning. More specifically, our model of the $\pi_{C,pi}$ may be

$$\log \pi_{C,pi} = \text{sum of the effects of the predictors and certain interactions,} \tag{2}$$

The right-hand side is a model as in analysis of variance, consisting of a constant, one-way margins ('main effects'), two-way margins ('two-way interactions'), etc. For the logarithmic transformation we have the same good reasons as in log-linear models. Firstly 'independence of probabilities' conveniently translates into additivity of the model for log probabilities. Secondly, contrasts such as the differences of two log probabilities are 'log odds' (=logits), which are easy to interpret, and thirdly the log transformation has several mathematical conveniences as well.

Note that for each p the sum $\Sigma_i \pi_{C,pi}$ must be one. If the right-hand side of (2) contains a parameter β_p, this can take care of the scaling. To this end the model must contain the main effect 'condition', a categorical variable with a different level for each condition. Moreover, we have a set of say q predictors, the values of which only depend on the condition p and the categories i of C. Let us denote these values by $x_{C,pik}$ (k=1..q), then our model is

$$\log \pi_{C,pi} = \beta_p + \Sigma_k x_{C,pik}\beta_{ik} \text{ (for all } p,i). \tag{3}$$

For condition p we define

$$f_{C,pi} = \text{the observed frequency of } C = i \text{ after condition } p,$$
$$n_p = \text{the observed frequency of condition } p \tag{4}$$

Thus the expected frequency $\mu_{C,pi} = Ef_{C,pi}$ equals $n_p\pi_{C,pi}$, and $\log \mu_{C,pi} = \log \pi_{C,pi} + \log n_p$. As $\log n_p$ can be absorbed into the term β_p in (3), we see that this model is equivalent to a similar linear model for log expected frequencies rather than log expected probabilities.

In our case we are only interested in modelling the marginal distributions $(C \mid p)$ and $(T \mid p)$, i.e. we assume that child and therapist behaviour are independent conditionally upon p. This implies that we assume there are no instantaneous effects, and that we have included all relevant predictors. This situation differs from 'univariate multinomial response models' in the same way as MANOVA differs from ANOVA, as we have now a 'multivariate multinomial response model'.

3.2. The distributional model

For the parallel streams of behaviour produced by our (fictitious) single child-therapist pair, the k-th order Markov chain assumption states that the successive realisations of the behaviours C and T conditioned upon C_{-1}, T_{-1}, ..., C_{-k+1}, and T_{-k+1} are independently and identically distributed. The frequency distribution obtained from independently and identically distributed observations with a finite sample space has a multinomial distribution. Information about the $\pi_{C,pi}$ comes from contingency tables like

$$C \times C_{-1} \times T_{-1} \times C_{-2} \times T_{-2}\ , \tag{5}$$

where C should be considered as the response variable and the others as predictors. A similar table can be defined for $\pi_{T,pj}$. The frequencies defined in (4) are the frequencies of this table. Thus for any fixed p the two times four frequencies $(f_{C,pi})$ and $(f_{T,pj})$, $i,j = 1..4$ are multinomially distributed, and for different p they are independent. Hence, the two sets have within and between themselves product-multinomial distributions. It should be mentioned that we do not need to have complete (=completely cross-classified) tables of predictors. In section 7 we will model the off-diagonal elements of the transition matrix $C \times C_{-1}$ without considering the diagonal elements.

The above development has shown that the Markov chain model naturally leads to a multinomial response model. Software for multinomial response models is SAS CATMOD (SAS

Institute Inc., 1985), MULTIQUAL (Bock & Yates, 1973), FREQ (Haberman, 1979), and SPSS LOGLINEAR (SPSS Inc., 1988).

4. Estimation and implementation

As estimators BMDP4F (Dixon et al., 1988) and SPSS offer the maximum likelihood estimator, SAS offers a choice between the maximum likelihood estimator and the weighted least squares estimator (WLS). The WLS option also allows the simultaneous analysis of several multinomial response variables.

There are two well-known algorithms for maximum likehood estimation, i.e. Newton-Raphson and Iterative Proportional Fitting (is equal to the Stephan-Deming algorithm). Therefore the particular algorithm used in a package need not concern the user. (BMDP4F, SPSS HILOGLINEAR use IPF; SPSS LOGLINEAR and SAS CATMOD use Newton-Raphson.) The WLS estimator is roughly equivalent to a one-step Newton-Raphson estimator (Grizzle, Starmer, Koch, 1969). One of the problems with this approach is that cells with random zeroes have to be replaced with a small positive number in order to be included in the analysis.

The SPSS programs and SAS CATMOD unfortunately do not produce all the parameter estimates, but only the 'independent ones'. For example the last category is not computed, and the manuals indicate that one can calculate them as 'minus the sum of the other estimates', which is correct, but one wonders why the computer programs do not do the calculations for us. Moreover, the computation of estimated standard errors for such omitted estimates is quite involved. The clumsy solution is to run the programs each time with different 'dependent' categories. This is computationally laborious, and quite expensive, especially for higher order interactions. BMDP does produce all estimates, but only handles log-linear and binomial logit models.

5. Model selection for child behaviour as response

For the substantive research both the child and therapist response functions were modelled separately. However, in this paper we will only do so for child behaviour. Model selection was performed with the IBM mainframe version of the program LOGLINEAR in SPSS^X, version 3.0 (SPSS Inc., 1988), and as a check parallel analyses were run with the mainframe SAS Version 5.0 program CATMOD (SAS Institute Inc., 1985). This yielded identical results. The simultaneous analyses with both therapist and child as multinomial response variables (see section 6) were carried out with SAS CATMOD.

5.1. Deviance-df plots

For model selection we use scatter plots (Deviance-df plots) in which each model is entered as a single point. The ordinate of the plot is the deviance (= -2 log likelihood) of the model (G^2), the abcissa is the degrees of freedom (df) of the model. The difficult part of model selection is the comparison of non-nested models. We use the following heuristic to simplify the situation. If model A has less df and more deviance than model B, we ignore model A. In the deviance-df plots this means that any model dominates all models North-West of it. Thus we only have to choose between models on the South-East boundary. Along this boundary we have a range of models from poorly fitting parsimonious models up to well-fitting complicated models. In general a proper balance between goodness-of-fit and simplicity should not be based on merely statistical grounds, but at least partially on substantive insights. In most cases along the boundary the ratios (deviance/df) and (change in deviance)/ (change in df) increase as a function of df. These can help us to choose a model, for example, by use of Bonnett and Bentler's (1983) index

$$\delta = 1 - \{(\text{deviance}_1/\text{df}_1)/(\text{deviance}_0/\text{df}_0\}, \qquad (6)$$

where the 0 indicates the null or comparison model, and the 1 the model to be compared. For a qualitative discussion of the stochastic properties of these plots, see Verbeek (1984).

5.2. Model search

The basis for the first analysis is a 4^5-table (C by C_{-1}, C_{-2}, T_{-1}, T_{-2}). A preliminary screening was carried out to establish an appropriate set of models. All models considered are displayed in the deviance-df plot (Figure 1). The most South-East (=dominating) models are connected with a line (the convex hull). In the present data, the set of dominating models form a hierarchical set, i.e. the models on the lower part of the curve contain all terms of the models on the higher part of the curve. To find an acceptable model within such a hierarchical set one may use Bonnett and Bentler's (1983) index for normed fit δ.

Table 1 gives an overview of a selection of the models which include one or more interaction terms with the response variable. The models in Table 1 are those with only the response variable C, which will be used as the null or comparison model (Model 1); Child behaviour at the previous measurement: C_{-1} (Model 2); Child behaviour at two measurements earlier added: $C_{-1}+C_{-2}$ (Model 3); Addition of therapist behaviour at t_{-1}: $C_{-1}+C_{-2}+T_{-1}$ (Model 4); Therapist behaviour at t_{-2} included as well: $C_{-1}+C_{-2}+T_{-1}+T_{-2}$ (Model 5). After all the one-way effects are included, the most important two-way interaction is that of the child at t_0 with the child and therapist at t_{-1}: $C_{-1}xT_{-1}+C_{-2}+T_{-2}$. By now the increase in δ becomes so small that including further terms does not substantially improve the descriptive adequacy as measured by δ. To choose an adequate

Table 1. *Models for Child Behavior at t_0*

Model[a]	G^2	df	δ	Remarks
1. I^b	71651	765	0.00	Null model
2. C_{-1}	11094	756	0.84	Child lag 1 only
3. $C_{-1} + C_{-2}$	4835	747	0.93	All child effects
4. $C_{-1} + C_{-2} + T_{-1}$	3217	738	0.95	** Preferred model **
5. $C_{-1} + C_{-2} + T_{-1} + T_{-2}$	2723	729	0.96	All main effects
6. $C_{-1}xT_{-1} + C_{-2}$	2218	711	0.97	
7. $C_{-1}xT_{-1} + C_{-2} + T_{-2}$	1802	702	0.97	
8. $C_{-1}xT_{-1} + C_{-1}xC_{-2} + T_{-2}$	1259	675	0.98	

Note:[a] Total number of units = 71276. [b] *I* indicates the model with only a constant, which is different for each logit of *C*.

model there exist few guidelines beyond settling for a δ somewhere in the .90's, and requiring the

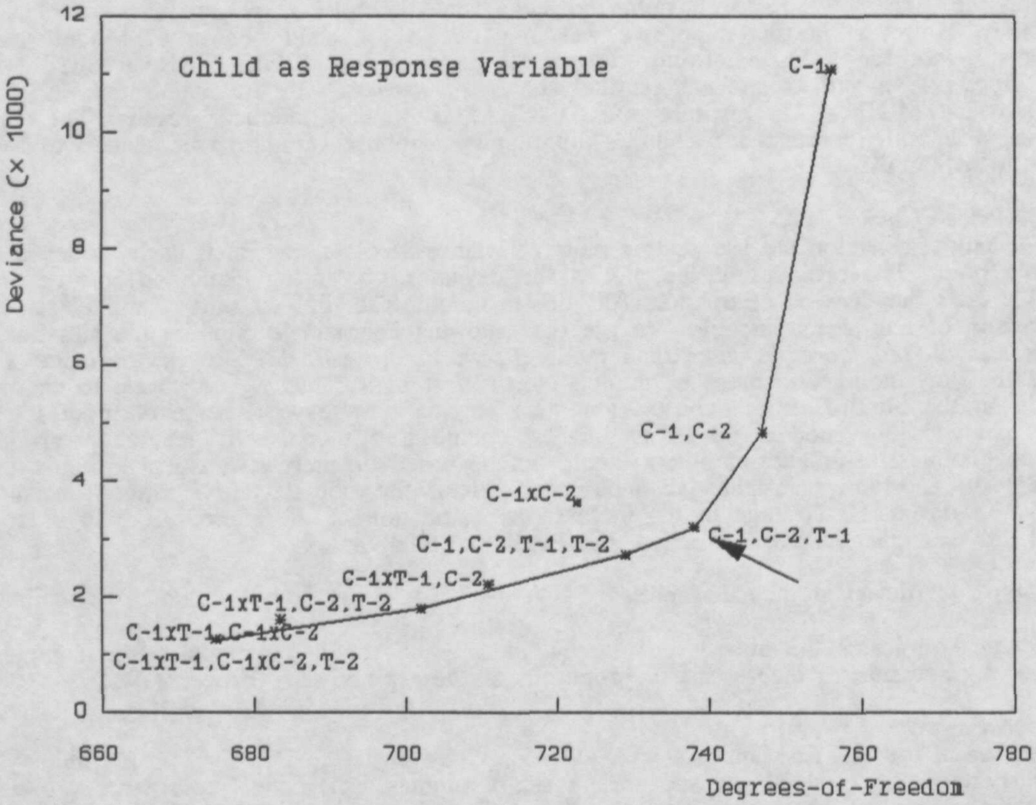

Figure 1 *Dev-Df Plot - Child as Response Variable*

last step to give some 'interesting' increase in the value of δ. Substantive arguments can play an important role. Considering the large number of observations, and the research interest in the behaviour of both the child and the therapist, one could settle for reporting the $C_{-1} + C_{-2} + T_{-1}$ model. The differences between the estimates for the parameters common to Model 4 and the next one (Model 5) were negligible for C_{-1} and C_{-2}. For T_{-1} the values were generally somewhat lower than for Model 4 but their pattern was essentially the same. The main reason for this seemed to be the similarity of the patterns of T_{-1} and T_{-2}, with T_{-1} having rather higher values than T_{-2}.

The 4^5-table ($C, C_{-1}, C_{-2}, C_{-3}, T_{-1}$) allows us to investigate the even longer range effect of the child behaviour. Even though 71,276 observations for 1024 cells seem ample, the present table has two zero entries in one of the four-dimensional margins, and many very small ones. The primary reason for this phenomenon is the general continuity of behaviour, which means that certain sequences of behaviour, especially four consecutive changes of category, are quite rare. Such zero cells may influence the stability of the parameter estimates, and they make the estimation of standard errors imprecise. Thus, some caution is in order when interpreting effects in models based on this table. During model selection, the first acceptable model was $C_{-1}+C_{-2}+C_{-3}$ with characteristics (G^2,df,δ) equal to (4209, 732, .94), taking precedence over the model of the previous section, $C_{-1}+C_{-2}+T_{-1}$, with (4903, 732, .93), but the difference in δ is hardly interesting. The next acceptable model is $C_{-1}+C_{-2}+C_{-3}+T_{-1}$ with (2938, 723, .95). This emphasizes the strong continuity in the behaviour of the child as the lag-3 behaviour has a slight edge over the lag-1 therapist behaviour. It also shows that to model child behaviour, one needs at least a third-order Markov chain.

6. Child and therapist behaviour as response

To evaluate the relative sizes of the autodependencies of, and crossdependencies between C and T, the most obvious 4^5-table, $CxTxC_{-1}xT_{-1}xC_{-2}$, already contains 107 sampling zeroes, therefore we decided to analyze the $CxTxC_{-1}xT_{-1}$-table using the WLS option of SAS CATMOD, with both C and T as response variables and C_{-1}, T_{-1}, and $C_{-1}xT_{-1}$ as predictor variables. Separate analyses for each of the dependent variables with SPSS^X LOGLINEAR would have resulted in the same values for the effects, but due to the lack of estimates for the covariances between the predictors no tests would have been available.

Table 2. *Multivariate Logits for C and T as Response Variables and T_{-1} and C_{-1} as Predictor Variables - Model: $T_{-1}xC_{-1}$*

Effect	Child Behavior at t_0: NonPlay vs ImagPlay	PrepPlay vs ImagPlay	FuncPlay vs ImagPlay	Therapist Behavior at t_0: NonPlay vs ImagPlay	PrepPlay vs ImagPlay	FuncPlay vs ImagPlay
Constant	.2	.9	.1	.7	-.3	-1.0
C_{-1}						
NonPlay	**1.6**	.6	.3	**.5**	.7	.4
PrepPlay	.4	**1.7**	.4	.4	**.9**	.2
FuncPlay	.6	.6	**2.3**	.2	.2	**1.3**
ImagPlay	**-2.6**	**-2.8**	**-3.0**	**-1.2**	**-1.9**	**1.9**
T_{-1}						
NonPlay	**.2**	*.1*	*-.1*	**.8**	*.1*	-.2
PrepPlay	.5	**.9**	.2	.4	**1.7**	.3
FuncPlay	.6	.4	**1.3**	.2	.3	**2.0**
ImagPlay	**-1.2**	**-1.4**	**-1.4**	**-1.4**	**-2.1**	**-2.1**
$C_{-1}xT_{-1}$ [*not displayed*]						

Notes: **Bold** values indicate the overall patterns of generally high values. *Italics* indicate non-significant main effects. All autodependencies and crossdependencies are pairwise significantly different, except those for PrepPlay at t_{-1} for both C_{-1} and T_{-1}. The boxed numbers are the crossdependencies and the other numbers the autodependencies.

In Table 2 we present the response functions for the log-odds with respect to ImagPlay, however the interactions between the predictors have been omitted from the table as we are not discussing their substantive interpretation anyway. Furthermore, to economize in space, the other three possible response functions not involving Imagery Play are not included either. The central question is whether the influence of the therapist on the child is larger than the reverse. In the present case this means testing the equality of twelve pairwise contrasts in the crossdependence blocks. Each pairwise contrast is significantly different, except those involving PrepPlay as the predictor. The table shows that when a significant difference occurs, the influence of the child on the therapist is mostly larger than vice versa. The minor exceptions occur for FuncPlay with the NonPlay/ImagPlay and PrepPlay/ImagPlay response functions, but the values are not overly large. Similar tests have been carried out for the autodependence effects. If effects differ significantly, they are always larger for the child than for the therapist. Indicating that child behaviour is more determined by own behaviour than therapist behaviour is. Child behaviour is more autonomous, and less influenced by the therapist, than vice versa.

7. New behaviour

In several aspects the analyses are only partly satisfactory from a substantive point of view. One primary question is the question of initiative. With the dominating continuity of behaviour (approximately 70%) it is difficult to see who initiates new behaviour, the child or the therapist. To look at this problem in more detail, one would have to look at only new behaviour at t_0. We will primarily focus on new behaviour of the child.

7.1. Analysing new behaviour with incomplete tables

Bishop, Fienberg, and Holland (1975; chap. 7) propose to analyse new behaviour by first transforming the time series by eliminating all repeated events. Thus AADABBBDDCCC would become ADABDC. As a consequence the diagonal of the first-order transition matrix will not be modelled. In second-order chains all sequences with two consecutive, identical codes become impossible e.g. ABB and BBD. The sequence ABBBD, originally yielding the triplets ABB, BBB, and BBD, becomes ABD and A four lags before D is supposed to be remembered by the child, and is modeled as second-order influence.

We propose to use (3) only for the non-diagonal cells of the child two-way transition matrix, but for second order chains, we will exclude triplets ending on the same codes, e.g. ABB and BBB. From the contingency table (5) we remove the diagonal plane of all cells (c, c_{-1}, t_{-1}, c_{-2}, t_{-2}) for which $c = c_{-1}$. Our actors have very short memories, the actors of Bishop et al. have long and variable length memories. Obviously either model could be right, depending upon the application. In this case, we prefer our model on theoretical grounds.

7.2. Computing response model estimates for incomplete tables

In SAS CATMOD WLS cannot handle the multicategory logit approach to multinomial response models for incomplete tables, and refers to ML. However, ML yields non-sensical estimates and standard errors in this case, and the df are incorrect. SPSS LOGLINEAR with logit contrasts also produces strange output and the same incorrect degrees of freedom. Apparently the combination of multicategory logits and incomplete tables has not been given appropriate attention in the design of these programs. In general users are ill-advised to try to use multivariate logits for incomplete tables. As some probabilities are missing, it is not immediately clear which pairs of categories of the dependent variable are used in such an analysis. If one specifies a multinomial response model for incomplete tables, the specification is unambiguous (see Verbeek and Kroonenberg, in preparation). In BDP4F there seems to be another kind of error, when one specifies an incomplete table. The df are in order, but we think there is an error in the rank of the model matrix.

A further complication is that due to the elimination of all cells with equal values for C and C_{-1} from the model, the corresponding parameters for the CxC_{-1} interaction should have been eliminated as well. However, neither SPSS LOGLINEAR nor SAS CATMOD seem to recognize this, and eliminate parameters in incorrect places. There are at least two ways to deal with this problem. The first is to construct the proper model matrix (or design matrix), but only SAS CATMOD has facilities for doing this. In our case this meant constructing a 192x92 model matrix, which we did with APL (see Verbeek & Kroonenberg, in preparation, for details). A second approach would be to eliminate the diagonal cells by adding additional predictors which cause the diagonal cells to be fitted perfectly. To this end one would have to add a predictor that has a different level for each combination of the condition p and the cells with equal c and c_{-1}. This solution could be used in both programs, but requires in our case a predictor with 64 levels or 63 dummy variables used as covariate(s).

Table 3. *Loglinear Models for New Child Behavior at t_0*
(with Non-Modelled Cells)

Model	G^2	df	δ	Remarks
1. I^a	5872	125	-	
2. C_{-1}	5774	120	0.00	Null model
3. $C_{-1} + T_{-1}$	2904	111	0.65	Also therapist lag 1
4. $C_{-1} + C_{-2}$	1855	111	0.46	Child lag 1 and 2
5. $C_{-1} + T_{-1} + C_{-2}$	396	102	0.92	** Preferred model **
6. $C_{-1}xT_{-1}$	2769	96	0.30	Lag 1 interaction

Note: [a] *I* indicates a model with only a constant, which is different for each logit of *C*.
Total number of observations is 20768.

7.3. Analysing the child's new behaviour

As before our analysis proceeds by first initiating a model search and subsequently estimating the parameters for the preferred model. Considering the arguments above, we performed a model search with SPSS^X LOGLINEAR, and out of methodological interest the parallel analyses were run with BMDP4F and SAS CATMOD. In Table 3 the results of this model search are presented. The maximum likelihood results with SAS were identical, but those using generalized least squares deviated substantially at several places. As mentioned above BMDP produced the same G^2 as the other programs, but failed to find satisfactory parameter estimates.

Unlike in the situation for complete tables, Model 1 in Table 3 cannot function as a null model of a hierarchy, because the effect of the not modelled cells on the number of estimates are not yet evident in this model, whereas this is the case in Model 2. Note, furthermore, that Models 2, 3, and 5 form a hierarchy, so do Models 2, 4, and 5, and Models 2, 3, and 6. In the last hierarchy it is evident that adding the $C_{-1}xT_{-1}$ interaction is a bad proposition, because δ decreases rather than increases. The preferred model is clearly Model 5, as δ has a reasonable value.

The way the log-linear analysis is specified in SAS the constant term always is estimated separately, so that one the CxC_{-1} interaction contains 16 - 1 (constant) - 4 (not-modelled cells) = 11 independent parameters, as can be seen in Table 4. As the parameter estimates include the row and column effects they do not add up to zero, and their values are not directly comparable to those of the other interactions.

Table 4 *Multivariate Loglinear Parameters for C as Response Variable and C_{-1}, C_{-2} and T_{-1} as Predictor Variables*

	Child Behavior at t_0			
Effect	NonPlay	PrepPlay	FuncPlay	ImagPlay
C_{-1}				
NonPlay	**x.x**	1.2	.7	-.1
PrepPlay	1.6	**x.x**	1.2	.5
FuncPlay	.7	1.1	**x.x**	-.4
ImagPlay	.3	.7	**x.x**	**x.x**
C_{-2}				
NonPlay	**.5**	-.1	-.2	-.2
PrepPlay	-.0	**.7**	-.3	-.4
FuncPlay	-.1	.3	**1.0**	1.2
ImagPlay	-.3	-.9	-.5	**1.7**
T_{-1}				
NonPlay	**.0**	.2	-.1	-.1
PrepPlay	.1	**.4**	-.0	-.5
FuncPlay	-.1	-.2	**.7**	-.4
ImagPlay	-.0	-.3	-.6	**1.0**

Note: The C_{-1} parameters are not directly comparable to the other parameters as they also contain the one-way margins.

The following conclusions may be drawn from the CxC_{-1} interaction. New NonPlay behavior can be twice as well predicted from PrepPlay, compared to FuncPlay and twice again as well as from ImagPlay. New PrepPlay can about equally well predicted from NonPlay as from FuncPlay, but worse from ImagPlay. New FuncPlay is twice as well predicted from PrepPlay as from NonPlay, and not from ImagPlay. Finally, New ImagPlay is only somewhat predictable from PrepPlay but NonPlay and especially FuncPlay are a counter indication for it to occur.

From the CxC_{-2} interaction, we see that the old pattern reestablishes itself again, each behavior is best predicted from the same behavior two time points ago. This probably indicates that longer chains of the same behavior are occasionally broken by single scores in other categories. The situation is similar for the CxT_{-1} interaction (see also Table 2). Several interesting details arise here, but they form part of the substantive discussion presented elsewhere.

8. Discussion

In this paper we have shown how multinomial response models may be fruitfully used to analyse parallel streams of behaviour, both when all behaviour needs to be modelled and when only new behaviour is of interest. Even though for this purpose standard statistical package can be used, employing them is certainly not a routine matter. As indicated several improvements could be made to facilitate analysing data such as ours, in particular presenting all parameters and their standard errors, and facilities to specify parameters to be deleted from the model matrix. Furthermore several errors in some of the programs need to be corrected.

References

Allison, P. D., & Liker, J. K. (1982). Analyzing sequential categorical data on dyadic interaction: A comment on Gottman. *Psychological Bulletin*, *91*, 393-403.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, Mass.: MIT Press.

Bock, R. D., & Yates, G. (1973). *MULTIQUAL: Log-linear analysis of nominal and ordinal qualitative data by method of maximum likelihood*. Chicago: International Educational Services.

Bonnett, D. G., & Bentler, P. M. (1983). Goodness-of-fit procedures for the evaluation and selection of log-linear models. *Psychological Bulletin*, *93*, 402-414.

Budescu, D. V. (1984). Tests of lagged dominance in dyadic interaction. *Psychological Bulletin*, *96*, 402-414.

Dixon, W. J. et al. (1988). *BMDP statistical software manual* (Vol. 1). Berkeley, CA: UCLA Press.

Grizzle, L. E., Starmer, C. F., & Koch, G. G. (1969). Analysis of categorical data by linear models. *Biometrics*, *25*, 489-504.

Haberman, S. J. (1979). *The analysis of qualitative data* (Vol.2). New York: Academic Press.

Harinck, F. J. H., & Hellendoorn, J. (1987). *Therapeutisch spel: Proces en interactie*. [Therapeutic play: Process and interaction]. Lisse, The Netherlands: Swets & Zeitlinger.

Hellendoorn, J., Kroonenberg, P. M., & Harinck, F. J. H. (1990)*Interaction in guided imaginative play*. Department of Education, Leiden University (submitted).

Iacobucci, D., & Wasserman, S. (1988). A general framework for the statistical analysis of sequential dyadic interaction data. *Psychological Bulletin*, *103*, 379-390.

SAS Institute Inc. (1985). *SAS user's guide: Statistics Version 5 edition*. Cary, NC: SAS Insitute Inc.

SPSS Inc. (1988). *SPSS-X user's guide* (3rd edition). Chicago, IL: SPSS Inc.

Verbeek, A. (1984). The geometry of model selection in regression. In T. K. Dijkstra (Ed.), *Misspecification analysis*. Berlin, FRG: Springer.

Verbeek, A., & Kroonenberg, P. M. (in preparation). *Multinomial response models, logit models, and log-linear models: Computational aspects.* Internal report. University of Utrecht.

Wasserman, S., & Iacobucci, D. (1987). Sequential social network data. *Psychometrika*, *53*, 261-282.