Categorical time series in psychological measurement
van Rijn, P.W.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
EPILOGUE

This thesis discussed methods for the analysis of psychological measurements in the form of categorical time series. The discourse started out with the analysis of multivariate continuous and categorical time series within the framework of structural equation modelling (SEM). The final part of the thesis concerned the analysis of multi-subject multivariate categorical time series by making reference to the frameworks of state space modelling (SSM) and item response theory (IRT). This final chapter consists of a short review of the general conclusions that can be drawn from the investigations in Chapters 2 to 5, i.e., the core of this thesis. Finally, guidelines for future research in the field of psychological measurement in the form of categorical time series are suggested.

6.1 Conclusions

6.1.1 Structural equation modelling

In Chapter 2, the use of the Toeplitz matrix containing sample auto- and cross-covariances in order to fit multivariate stationary autoregressive (AR) models was investigated. For normally distributed time series, it was found that parameters and standard errors are correctly recovered only for pure vector autoregressions. For multiple indicator autoregressive models, the estimated parameters were correct, but the standard errors were not. This limits the use of the Toeplitz method for this type of measurements. In the second part of the chapter, autoregressive models were fitted to multivariate categorical time series by using the Toeplitz matrix containing polychoric auto- and cross-correlations. The results indicated that the estimates of the parameters are correct for the used models, but the standard errors are not. In view of the simulation results and available asymptotic results on sample auto- and cross-covariances, it can be argued that the Toeplitz method should not be used in investigations with real data in which the type and order of the model is not known. Also, since stationarity is a necessary assumption for the Toeplitz method, but not for filtering methods, the latter are preferable in modelling time series measurements. The Toeplitz method can how-
ever have practical utility when used as a method of moments to obtain parameter estimates.

The above argumentation should not be taken as a case against applying the SEM framework in a time series setting (see also, MacCallum & Ashby, 1986). Rather, it is the methodology of using summary statistics for estimating and fitting models, a regular course of things in standard SEM applications, that is not recommended in the situation of normal and, especially, categorical time series.

### 6.1.2 State space modelling

Since filtering and smoothing methods to analyse categorical time series are not widely available, the second part of this thesis consisting of Chapters 3 to 5 focused on such methods. Specifically, the SSM framework was referred to for specifying and estimating models for categorical time series. Chapter 3 addressed the case of univariate categorical time series, whereas Chapters 4 and 5 discussed the multivariate case.

In Chapter 3, a Kalman filtering and smoothing method was used to fit a latent autoregressive process to an observed univariate categorical time series through a logistic response function. It was found that autoregressive parameters showed some bias, but could be consistently estimated. The estimates of threshold parameters, associated with particular categories of the time series, were however found to be biased and inconsistent. In general, the presented approach did not seem to work satisfactorily for univariate categorical time series.

The first part of Chapter 4 presented a case in favor of the analysis of intra-individual variation. The second part contained illustrations of a dynamic logistic model for multivariate dichotomous time series with simulated and real data. In this chapter, the presented dynamic logistic model was related to the Rasch model.

Chapter 5 elaborated on Chapter 4, with the focus shifting towards a comparison of the SSM and IRT frameworks. A comparison was made between standard IRT estimation methods and a Kalman filtering and smoothing method for the estimation of item and person parameters in the situation of cross-sectional data. The modelling approach presented in Chapter 4 was extended to polytomous time series. The chapter ended with an application of the methods to multiple subject polytomous time series. The presented approach provided a useful methodology for simultaneously fitting models to polytomous time series of multiple subjects.
The fact that the SSM framework is very useful for modelling time series measurements is not new. That it provides a meaningful framework for psychological time series measurements is therefore to be expected. In the present thesis, we have discussed models and estimation methods for categorical time series within the SSM framework, but also within the frameworks of SEM and IRT. That the SSM framework can be used in a sensible manner in combination with the familiar and comprehensive frameworks of SEM and IRT is one of the main conclusions of the present thesis. Particularly, the frameworks of SSM and IRT form a good combination for the type of categorical time series measurements which were analysed in this thesis. More general, the combination of these two frameworks can provide a unified approach to model cross-sectional, longitudinal, and time series data formed by categorical psychological measurements.

6.2 Guidelines for future research

In closing this thesis, guidelines for future investigations in psychological measurement in the form of categorical time series are suggested, which are distinguished in three interesting topics. First, further comparisons of the Kalman filtering and smoothing estimation method used in this thesis with other filtering methods are important. Many methods for non-normal time series and non-linear time series model are Bayesian (Doucet, de Freitas, & Gordon, 2001), and are known to work well. It would be of interest to compare Bayesian filtering methods with the method used in this thesis. Recently, Chow, Ferrer, and Nesselroade (2007) used an unscented Kalman filter to fit nonlinear dynamic models for dyadic interactions in emotions. Such a filter might also be applied to dynamic logistic models for categorical time series. It is evident that comparative studies are necessary to expose strengths and weaknesses of different methods.

A second important topic is the further development of parameter estimation methods apart from methods for filtering. Particularly, the estimation methods for parameters such as latent autoregressive parameters and variances of innovations are not well developed (Fahrmeir & Wagenpfeil, 1997). The numerical method used in the present thesis can then, for instance, be used for comparison.

Apart from the establishment of the performance of the estimation methods, there is a need to develop and investigate the quality of measures of fit for the particular time series models under scrutiny. This has clearly been a neglected issue in the present thesis, but is largely uncharted and remains one of the most important topics in future investigations. This because a crucial stage in any
statistical analysis is the selection of an appropriate model. Of special interest would be statistical fit procedures to assess if an observed psychological time series is or is not obtained from a stationary stochastic process. If nonstationarity is then found, an immediate need arises for developing statistical techniques for the fit of state space models with time-varying parameters (for the linear Gaussian model, see e.g., Molenaar, 1994). Whereas it is often straightforward to estimate complex statistical models, it is often more difficult to select the best of several straightforward statistical models.