

# The Development of Deviant and Delinquent Behavior of Adolescents: Applications of Latent Class Growth Curves and Growth Mixture Models

Jost Reinecke<sup>1</sup>

## Abstract

The article presents applications of different growth mixture models considering unobserved heterogeneity within the framework of *Mplus* (Muthén and Muthén, 2001, 2004). Latent class growth mixture models are discussed under special consideration of count variables which can be incorporated into the mixture models via the Poisson and the zero-inflated Poisson model. Four-wave panel data from a German criminological youth study (Boers et al., 2002) is used for the model analyses. Three classes can be obtained from the data: Adolescents with almost no deviant and delinquent activities, a medium proportion of adolescents with a low increase of delinquency and a small number with a larger growth starting on a higher level. The best model fits are obtained with the zero-inflated Poisson model. Linear growth specifications are almost sufficient. The conditional application of the mixture models includes gender and educational level of the schools as time-independent predictors which are able to explain a large proportion of the latent class distribution. The stepwise procedure from latent class growth analysis to growth mixture modeling is feasible for longitudinal analyses where individual growth trajectories are heterogeneous even when the dependent variable under study cannot be treated as a continuous variable.

## 1 Introduction

Longitudinal research studies with repeated measurements are quite often used to examine processes of stability and change in individuals or groups. With panel data it is possible to investigate intraindividual development of substantive variables across time as well as interindividual differences and similarities in change patterns. While the traditional analysis of variance (ANOVA) and the analysis of covariance (ANCOVA) assume homogeneity of the underlying covariance matrix across the levels of the between-subjects factors and the same covariance patterns for the repeated measurements, the structural equation methodology offers an alternative strategy: the *latent growth curve models*. These models describe not only a single individual's developmental trajectory, but also capture individual

---

<sup>1</sup> Faculty of Sociology, University of Bielefeld, Postbox 100131, D-33501 Bielefeld; jost.reinecke@uni-bielefeld.de

differences in the intercept and slopes of those trajectories. Based on the formative work of Rao and Tucker's basic model of growth curves (Rao, 1958; Tucker, 1958), Meredith and Tisak (1990) discussed and formalized the model within the structural equation framework. Further developments of the growth curve model were proposed by McArdle and Epstein (1987), McArdle (1988), Muthén (1991, 1997) and Muthén and Curran (1997).

The formal representation of a growth curve model can be seen either as a multilevel, random-effects model (Liang and Zeger, 1986) or as a latent variable model, where the random effects are latent variables (cf. for example Curran and Hussong, 2002: 69). In the latent variable model the repeated measurements are the manifest variables while the intercept and the slope are the latent variables (cf. the discussion in Hox, 2002, Chapter 14). Different specifications of growth curve models with structural equations using the programs LISREL (Jöreskog, K. G. and Sörbom, 2004) and EQS (Bentler, 2001) are discussed by Duncan, Duncan, Strycker, Li and Alpert (1999).

Observed heterogeneity in growth curve models can be captured by covariates (e. g. gender) explaining part of the variances of the intercept and slope. But the assumption of a single population underlying the growth curve has to be relaxed in the case of unobserved heterogeneity where different classes of individuals vary around different mean growth curves. A very suitable framework to handle the issue of unobserved heterogeneity is *growth mixture modeling* introduced by Muthén and Shedden (1999). These mixture models differ between continuous and categorical latent variables. The categorical latent variables represent mixtures of subpopulations where the membership to those subpopulations is inferred from the data. Like the conventional growth curve models, intercept and slope variables capture the continuous part of the model. The framework of growth mixture models can also be seen as an extension of the structural modeling approach with techniques of latent class analysis. The inferred membership of each individual to a certain class is produced with the information of the estimated latent class probabilities. Further developments and applications with the program *Mplus* (Muthén and Muthén, 2004) are discussed in several papers by Muthén (2001a, 2001b, 2003, 2004).

The simplest specification of a growth mixture model is the so-called *latent class growth analysis* where no variation across individuals are allowed within classes. This model was also discussed by Nagin and Land (1993), Nagin (1999) and Roeder, Lynch and Nagin (1999) with normal and non-normal outcomes especially focused on deviant and delinquent behavior (for a detailed overview see Nagin, 2005). Due to the substantive similarity of the data used in our study, the underlying distributions of the longitudinal measurements will be considered. Our measurements represent numbers of different deviant and delinquent behaviors and can be treated as count variables. The underlying statistical model of a count variable is the Poisson distribution (see, e. g., Ross, 1993). If the count variables are biased to zero, i. e., the particular behaviors seldom occur, the zero-inflated Poisson model (Lambert, 1992) should be the best statistical representation. A brief introduction of the growth mixture model including its special cases will be given in Section 2.

Major methodological developments in criminological longitudinal research are influenced by the debate whether distinctive groups about criminal behavior can be explored and in which way the development of a "criminal career" can be incorporated in a statistical model. The major controversy stems from fundamental disagreements about

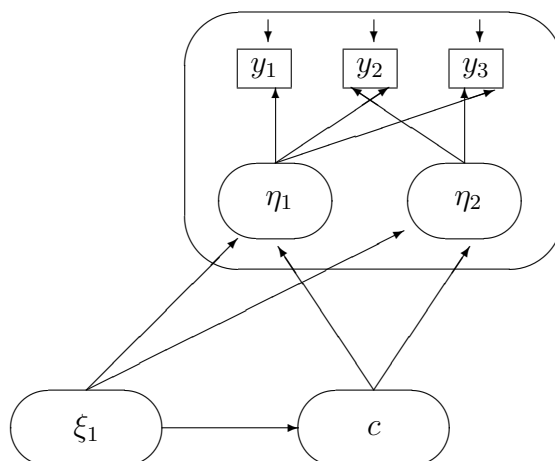
whether people with criminal activities form a distinctive group and in which way those "careers" can be explored (Nagin and Land, 1993). Several long-term studies, like the "Cambridge Study" (Farrington and West, 1990), the "Philadelphia Study" (Tracy, Wolfgang and Figlio, 1990) or the "Montreal Study" (Tremblay, Desmarais-Gervais, Gagnon and Charlebois, 1987) were analyzed by Nagin and collaborators to find population heterogeneity in behavioral trajectories. Depending on the type of the dependent variable, the nature of the sample and the characteristics of the community, three to five trajectories reflecting different intensity and growth of delinquency were detected. These analyses can differ between non-offenders, a time-limited delinquent behavior through adolescence and a more or less chronic group of offenders (D'Unger, Land, McCall and Nagin, 1998; Nagin 1999). Furthermore, background variables like household income or educational level of the parents are included via a multinomial logit model to explain differences in growth of criminal offending (Land, McCall and Nagin, 1996; Nagin, 1999). The major statistical limitation of those analyses is the assumption that within each group the individual developments are the same meaning that the growth parameters have no variance. This restriction can be relaxed within the growth mixture model of Muthén and collaborators. Although there are several methodological and statistical discussions about growth mixture models (e. g., Bauer and Curran, 2003, 2004), the systematic evaluation of these models with empirical data are still rare.

Applications of latent class growth models and growth mixture models within the framework of *Mplus* are the main focus of the present study. A representative panel study of adolescents' deviant and delinquent behavior will serve as data base for the systematic study of growth mixtures. After an introduction of the mixture models and their special cases (Section 2), the design of the study, the variables and descriptive statistics are introduced (Section 3). Section 4 will address the number of distinctive groups in the data under the assumption of the latent class growth model and explore substantive differences between those groups. Gender and level of education will serve as time-invariant predictors in the conditional models. Based on these results, Section 5 studies various growth mixture models allowing random intercepts and slopes. Differences to the previous results of the latent class growth analyses will be emphasized. Section 6 summarizes the results and discuss suggestions for further research with growth mixture models.

## 2 Method and models

### 2.1 Growth mixture models

The possibility that the individual trajectories of a dependent variable can vary is one of the main advantages of the growth curve model. This variation is captured by the growth curve factors. But the model assumes that all individuals are drawn from the same population, i. e., the means of the random effects have no variance. Growth mixture modeling relaxes this assumption and gives information about parameter differences across unobserved subpopulations. Instead of considering individual variation of a single mean of the intercept and slope, the growth mixture model allows different classes of individuals to vary around different intercepts and slopes (Muthén and Shedden, 1999). The classes are introduced by a latent categorical variable where the categories (classes) represent the



**Figure 1:** General growth mixture model.

unobserved heterogeneity of the data.

Following Muthén (2002, 2004) a growth mixture model (GMM) contains both latent growth variables  $\eta$  and a latent categorical variable  $c$  for  $c = 1, 2, \dots, K$ . The variable  $c$  captures latent trajectory classes representing the different subpopulations. An unconditional growth mixture model can be formalized by the following equation:<sup>2</sup>

$$\begin{aligned} y_{tk} &= \lambda_{1tk}\eta_{1k} + \lambda_{2tk}\eta_{2k} + \epsilon_{tk} \\ \eta_{1k} &= \alpha_{1k} + \zeta_{1k} \\ \eta_{2k} &= \alpha_{2k} + \zeta_{2k} \end{aligned} \quad (2.1)$$

$y_{tk}$  is the manifest variable measured in wave  $t$  and class  $k$ ,  $\eta_{1k}$  is the *initial level factor* whereas  $\eta_{2k}$  is the *linear growth factor*. To estimate the model, factor loadings  $\lambda_{1tk}$  are restricted to one for all  $y_{tk}$  whereas factor loadings  $\lambda_{2tk}$  are usually restricted to represent linear growth:  $\lambda_{2tk} = 0, 1, 2, \dots, T - 1$  for  $t = 1, \dots, T$ .  $\epsilon_{tk}$  is the random error term. In an unconditional model the latent variables are only described by their class specific means ( $\alpha_{1k}, \alpha_{2k}$ ) and variances ( $\zeta_{1k}, \zeta_{2k}$ ).

For a conditional growth mixture model the structural equations of 2.1 are extended by exogenous latent variables  $\xi_n$ :

$$\eta_{mk} = A_k + \Gamma_k \xi_{nk} + \zeta_{mk} \quad (2.2)$$

Matrix  $A_k$  contains the levels and slopes within  $k$ -classes while matrix  $\Gamma_k$  refers to the regressions of  $\xi_n$  within the  $k$ th class. Figure 1 gives an example of a growth mixture model with one latent exogenous variable  $\xi_1$  (a measurement model for the exogenous variable  $\xi_1$  is omitted in the figure as well as in the description of the model).

The relation between the exogenous variable  $\xi_1$  and the categorical class variable  $c$  is given by a multinomial logistic regression model:

$$\text{logit}(\pi_k) = \alpha_c + \Gamma_c \xi_1 \quad (2.3)$$

<sup>2</sup>For simplicity, person suffix  $i$  is omitted in the following equations.

with  $\pi_k = P(c_k = k|\xi_1)$  and  $\alpha_c$  as a  $(K - 1)$ -dimensional parameter vector.  $\Gamma_c$  is a  $(K - 1) \times q$ -parameter matrix containing regression coefficients of  $K$  classes on predictor  $\xi_1$ . Assuming two classes, matrix  $\Gamma_c$  contains one parameter for the first class while the second parameter stands for the reference class. In *Mplus*, the reference class is always the last class of the mixture model. With an unordered number of categories  $k$  of the latent class variable, the probability expression  $P(c_k = k|\xi_1)$  can be formalized as follows (Muthén and Muthén, 2004: 346):

$$P(c_k = k|\xi_1) = \frac{e^{\alpha_{c_k} + \gamma'_{c_k} \xi_1}}{\sum_{k=1}^K e^{\alpha_{c_k} + \gamma'_{c_k} \xi_1}} \quad (2.4)$$

For the last class  $e^{\alpha_{c_K} + \gamma'_{c_K} \xi_1}$  equals 1. Therefore, the probability to be in class  $c$  is simply the ratio of the exponentiated value of the logit and the sum of all exponentiated values. The categorial part of the growth mixture model in Figure 1 can be enlarged by adding *outcome variables* or *distal indicators* which are formally incorporated as a logistic regression with covariates  $c$  and  $\xi_1$  (cf. Muthén, 2004: 349). The applications in Sections 4 and 5 will not consider this part of the mixture model.

## 2.2 Model estimation and evaluation

Growth mixture models are estimated by maximizing the log likelihood function within the admissible range of parameter values given classes and data. *Mplus* (Version 3.11) uses the principle of maximum likelihood estimation and employs the EM-algorithm for maximization (Dempster, Laird and Rubin, 1977; Muthén and Shedden, 1999). A set of starting values are provided by the program before one set is used for the final estimation of the parameters. Version 3 of *Mplus* includes a so-called *integration method* which tests sets of different starting values evaluating the maximum initial stage log likelihood value. The seed number corresponding to that value is used for the final estimation of the model. For re-estimation of the model parameters the *optimal seed value* of the previous run can be included in the input file (for details see Muthén and Muthén, 2004). For a given solution, each individual's probability of membership in each class is estimated. Individuals can be assigned to the classes by calculating the posterior probability that an individual  $i$  belongs to a given class  $k$ . Each individual's posterior probability estimate for each class is computed as a function of the parameter estimates and the values of the observed data (cf. Muthén and Muthén, 2001: 367f.).

It is always an empirical question how many classes are sufficient to describe the unobserved heterogeneity of the data. By classifying each individual into his most likely class, a table with rows corresponding to individuals classified into a given class can be constructed. The columns of that table show the average conditional probabilities to be in the particular class. High diagonal and low off-diagonal probabilities indicate a good classification. The entropy measure  $E_K$  summarizes the quality of the classification (Muthén and Muthén, 2001: 372):

$$E_K = 1 - \frac{\sum_i \sum_k (-\hat{p}_{ik} \ln \hat{p}_{ik})}{n \ln K} \quad (2.5)$$

$\hat{p}_{ik}$  denotes the estimated conditional probability for individual  $i$  in class  $k$ .  $E_K$  ranges from zero to one, where values close to one indicate a good classification of the data.

Maximum likelihood estimates and standard errors are obtained by maximizing the log likelihood function, given the observed data and the number of classes. Standard errors of estimates are asymptotically correct if the underlying mixture model is the true model. But in mixture models a  $k$  class model is not nested within a  $k + 1$  group model. Therefore, conventional mixture tests like the Akaike Information Criterion (AIC; Akaike, 1987) and the Bayesian Information Criterion (BIC; Schwarz, 1978) have to be used for model comparisons.

In general, test statistics require well-defined classes in a mixture model. The likelihood ratio comparing a  $k - 1$  and a  $k$ -class model does not have the usual large-sample  $\chi^2$ -distribution because the class probability parameter can be at the boundary of its admissible space, i. e., a latent class probability can be zero (Rudas et al., 1994). In addition to that, a  $\chi^2$ -difference between two models is only suitable for model selection when the models are nested. In mixture models a  $k$  class model is not nested within a  $k + 1$  group model. Therefore, the BIC is being used for model comparisons:

$$BIC = -2\log L + pln(n) \quad (2.6)$$

$L$  is the value of the model's maximized likelihood,  $p$  is the number of parameters and  $n$  the sample size. The second term in both equations is the so-called *penalty term* which penalize an increase of the likelihood with additional parameters. In addition, the BIC includes the sample size. Usually, the model with the smallest BIC is accepted within model comparisons. Furthermore, *Mplus* calculates a sample size adjusted BIC with  $n = (n + 2)/24$  which was found to give superior performance for model selection (Yang, 1998). If the  $k$ -class model contains a redundant class, the  $k - 1$ -class model with the smaller BIC value should be chosen. By adding a class an expansion of the model is only desirable if the resulting improvement in the log likelihood exceeds the penalty for more parameters. But to accept or reject a model on the basis of the BIC is more or less descriptive and does not imply any statistical test.

Lo, Mendell, and Rubin (2001) proposed a likelihood ratio-based method for testing  $k - 1$  classes against  $k$  classes in mixture models. The Lo-Mendell-Rubin likelihood ratio test (LMR-LRT) considers the usual likelihood ratio for testing the  $k - 1$  model against a  $k$  model but with the correct distribution. The  $p$ -value from the test represents the probability that  $H_0$  is true, i. e., that the model is sufficient with one less class. Therefore, a low  $p$ -value indicates that the  $k - 1$  class model has to be rejected and the  $k$ -class model can be accepted for substantive interpretations. LMR-LRT has been criticized by Jeffries (2003), but importance of the critics in applications is unknown (Muthén, 2004: 356). BIC, adjusted BIC and the LMR-LRT test will be used for model selection in Sections 4 and 5.

## 2.3 Special cases of growth mixture models

### 2.3.1 Latent class growth analysis

A special case of the growth mixture model is the latent class growth analysis (LCGA), which has been studied by Nagin and Land (1993), Nagin (1999) and Roeder, Lynch and Nagin (1999). Jones, Nagin and Roeder (2001) discuss different applications of latent

class growth models assuming Poisson distributed count data as well as the more general zero-inflated Poisson model discussed below. LCGA is a submodel of Equation 2.1 and characterized by having zero variances and covariances of the intercept and slope variables  $\eta$ . In an unconditional model the structural part of Equation 2.1 reduces to:

$$\begin{aligned}\eta_{1k} &= \alpha_{1k} \\ \eta_{2k} &= \alpha_{2k}\end{aligned}\tag{2.7}$$

Individuals within a class are treated as homogeneous with respect to their development. As Muthén (2004: 350) points out, LCGA has two major advantages: It can be used to find cut points in the within-class variation on the growth factors. This leads to the question which different latent classes represent substantially different trajectories and which classes exist only due to minor variations. On the other hand the latent classes can be viewed as a nonparametric representation of the distribution of the growth factors, resulting in a semi-parametric model (Nagin and Land, 1993; Nagin, 1999). LCGA serves here as a starting point for the growth mixture model (GMM) with random intercepts and/or random slopes. The random component in the GMM is represented by a mixture of  $k$  classes. The stepwise procedure from LCGA to GMM will be discussed with substantive applications in Section 4 and 5.

To study development of deviant and delinquent behavior is one of the main topics in criminal sociology or criminology. Very often the longitudinal data gives information about the incidence rate of that behavior or the number of convictions. From a methodological point of view the distributions of those variables are counts and have to be treated differently compared to continuous data. The so-called "key approach in the modeling of delinquent and criminal careers" (Land et al, 1996) is the Poisson distribution with the corresponding regression models.

### 2.3.2 Poisson model and zero-inflated Poisson model

The Poisson model assumes count variables instead of continuous variables. Let  $Y = 0, 1, 2, \dots$  be a random variable for a given time interval and  $y$  be the number of observed occurrences. The number of events in an interval of a given length is Poisson distributed with the following probability density function:

$$Pr(Y = y) = e^{-\nu} \left[ \frac{\nu^y}{y!} \right]\tag{2.8}$$

The expected value or mean of the Poisson distribution is  $E(Y) = \nu$  with  $Var(Y) = \nu$ .<sup>3</sup> Usually, the parameter  $\nu$  is referred as the mean rate of occurrence of events. Small values of  $\nu$  yield high probability for zero occurrences of the random variable  $Y$ . The higher the value of  $\nu$ , the lower the skewness of the distribution. The Poisson distribution is considered most appropriate for modeling events which seldom occurs (for an overview with applications to criminal careers data, cf. Land et al., 1996). If  $\nu$  is indexed to each

<sup>3</sup>Note, that in the literature the greek letter  $\lambda$  is often used for the expected value of a Poisson distribution. In structural equation models  $\lambda$  is reserved for vectors of factor loadings. Therefore, I use the letter  $\nu$  here for the expected value.

individual  $i$  in a sample, exogenous variables  $X$  can be incorporated to specify a Poisson regression model (Land et al., 1996: 395):

$$\ln(\nu_i) = X_i\beta \quad (2.9)$$

Instead of a simple regression model the growth curve model can be used to explain the Poisson distributed count data:

$$\ln(\nu_{it}) = \lambda_{1t}\eta_1 + \lambda_{2t}\eta_2 + \epsilon_t \quad (2.10)$$

$\nu_{it}$  is the expected number of occurrences of the measurement  $y$  of individual  $i$  at time  $t$ .  $\eta_1$  is the intercept and  $\eta_2$  is the slope variable. Factor loadings  $\lambda_{1t}$  and  $\lambda_{2t}$  are restricted for linear growth equal to Equation 2.1. Equation 2.10 assumes that the growth parameters do not differ between unobserved groups.

To cover the unobserved heterogeneity, a Poisson-based latent class growth model can be formulated (Nagin and Land, 1993: 335; Nagin, 1999: 144; for specification in *Mplus* see Muthén and Muthén, 2004: 190):

$$\ln(\nu_{it}^k) = \lambda_{1t}^k\eta_1^k + \lambda_{2t}^k\eta_2^k + \epsilon_t^k \quad (2.11)$$

$\nu_{it}^k$  is the expected number of occurrences of the measurement  $y$  of individual  $i$  at time  $t$  given the membership in class  $k$ . The conditional number of events,  $P(y_{it}^k|k)$ , should follow the Poisson distribution.

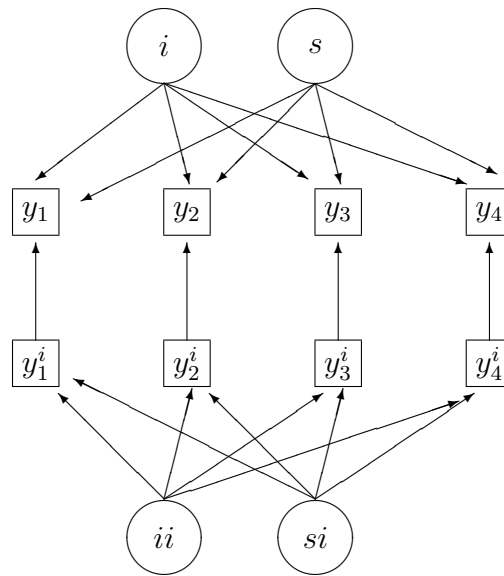
If the number of zeros in the count variable are very large, a variant of the Poisson regression model is more appropriate: the so-called *zero-inflated Poisson model* (ZIP) originally proposed by Lambert (1992). The ZIP model combines the regression model in Equation 2.9 with a logit model to cover the zero inflation in the count variable  $Y$  with probability  $p$  that  $Y$  is zero (Lambert, 1992: 3). Two parallel growth mixture models are estimated simultaneously when zero inflation of the data is assumed: The first model contains the count part of the outcomes with values of zero and above (Variables  $y_1$  to  $y_4$  in Figure 2). Intercepts of the outcomes are fixed to zero as the default. The means of the growth curve variables (Variables  $i$ ,  $s$ ) are estimated for each class. The second model refers to the zero-inflation part of the outcome with only values of zero in all measurements (Variables  $y_1^i$  to  $y_4^i$ ). Intercepts of the outcomes are estimated and held equal as the default. The mean of the intercept variable (Variable  $ii$ ) are fixed to zero for all classes while the mean of the slope (Variable  $si$ ) is estimated and held equal for all classes (cf. Muthén and Muthén, 2004: 190). Those two parallel growth curve models can be combined in *Mplus* to get simultaneous parameter estimates. If rare events (e. g. counts of deviant and delinquent behavior) are under study, the ZIP model should be more appropriate than the Poisson model (see the applications with PROC Traj in Jones et al., 2001).

### 3 Design of the study and descriptive statistics

The empirical basis for the following analyses with mixture models is taken from the longitudinal research project *Juvenile Delinquency in Modern Towns*.<sup>4</sup> The main focus of

<sup>4</sup>This interdisciplinary research project is located at the Universities of Münster (Institute of Criminology) and Bielefeld (Faculty of Sociology) and supported by the German National Science Foundation (DFG) under grant numbers Bo1234/6 and Re832/4.





**Figure 2:** Two-part growth mixture model with zero-inflated measurements.

the study is on the emergence and the development of deviant and delinquent behavior of juveniles and the social control surrounding it; both formal, meaning the police and the judiciary, and informal, referring to the school and the family (Boers et al., 2002).

The panel data contains self-administered interviews with pupils from the town Münster located north of Cologne in West Germany. The initial survey was conducted in the year 2000 with pupils from 7th, 9th and 11th grade considering all relevant school types. The school system in the community differ between three educational levels depending on pupils' achievements and their learning abilities. Further on, these educational levels will be labeled as *high*, *medium* and *low* in the analyses of the conditional models in Section 4 and 5. One cohort (7th grade) have been interviewed annually until the pupils reached the 10th grade in the year 2003. 31 schools participated over the whole time period. The mean response rate was about 85%. The relatively high response rate is due to confidentiality restrictions. Almost every longitudinal study with sensitive topics avoid to collect data which can identify respondents very easily (e. g. home addresses). Panel attritions can rise up to 50% depending on the time period under study. Instead of collecting pupils' home addresses, an individual encryption code was created. To generate the code, a specific code sheet were administered prior to the questionnaire in each panel wave. For example, the code sheet includes questions about specific letters of respondent' s eye and hair colors and numbers of respondent's birthday. Additional informations (e. g. change of class, school) were used to avoid equal codes belonging to different persons (see Pöge, 2005 for details).

Table 1 shows the distributions of gender and educational level for each panel wave. Compared to official statistics the distribution of the educational level is biased. Low educated pupils are somewhat underrepresented while the high educated pupils are overrepresented in the data. It should also be noticed that the capability to remember the answers to the code sheet questions generating the encryption code is correlated to the educational level of the schools. There is also a somewhat higher proportion of females

**Table 1:** Sample of respondents in the four-wave panel study (2000-2003).

|                    | Wave 1     | Wave 2     | Wave 3     | Wave 4     | Wave 1-4   |
|--------------------|------------|------------|------------|------------|------------|
| Gender             | n (%)      | n (%)      | n (%)      | n (%)      | n (%)      |
| male               | 665 (48.1) | 816 (48.9) | 868 (50.3) | 716 (48.1) | 353 (43.4) |
| female             | 717 (51.9) | 852 (51.1) | 857 (49.7) | 774 (51.9) | 460 (56.6) |
| $\Sigma$           | 1382 (100) | 1668 (100) | 1725 (100) | 1441 (100) | 813 (100)  |
| Level of Education | n (%)      | n (%)      | n (%)      | n (%)      | n (%)      |
| low                | 282 (20.4) | 402 (24.1) | 423 (24.5) | 341 (22.9) | 129 (15.9) |
| medium             | 410 (29.7) | 525 (31.5) | 560 (32.5) | 487 (32.7) | 241 (29.6) |
| high               | 690 (49.9) | 741 (44.4) | 742 (43.0) | 662 (44.4) | 443 (54.5) |
| $\Sigma$           | 1382 (100) | 1668 (100) | 1725 (100) | 1441 (100) | 813 (100)  |

(56%) in the four-wave panel data reflecting the greater ability to remember the answers of the code variables. Due to improvements of the encryption code sheet in the following panel waves, the bias could be reduced. In addition, a higher mobility between the schools and dropouts of 9th grade pupils after the third panel wave have to be considered. The last column of Table 1 presents the distribution of respondents that participated in all panel waves and could be detected via the encryption code (N=813). Those respondents are used for the mixture models described in subsequent sections.

Different types of deviant and delinquent behaviors (e. g., robbery, aggravated assaults, shoplifting) are asked in each panel wave for the respective last year. Only those types of delinquencies are considered which were asked in all four waves. In the average, most of them increase until the third wave (9th grade pupils) and decrease one year later. This development would give empirical evidence of an adolescent-limited group of offenders which were detected in other longitudinal studies (cf. D'Unger et al., 1998; Nagin, 1999). Using or dealing with drugs shows a somewhat different picture. The rates increases from 5% in the first wave up to 27% in the fourth wave for drug use and from nearly zero up to 4% for drug dealing in the fourth wave.

Table 2 shows the wave-specific overall prevalence rates as additive indexes of the particular delinquencies. The mean rate of deviant and delinquent behavior increased from 0.46 (7th grade) to the peak of 0.86 (9th grade). One year later the mean rate decreased to 0.82 (10th grade). The percentage of zeros reflecting no particular delinquency of the pupils decreased from 77% (7th grade) to nearly 63% (10th grade).

Correlations within the overall-prevalence rates reflect a time-dependent pattern (Table 3). The longer the time distance between the measurements, the lower the correlations. In accordance with our expectations both exogenous variables correlate negative to the prevalence rates: Females show less deviant and delinquent behavior than males, pupils on schools with a higher educational level are also less deviant and delinquent than pupils from schools with a lower educational level. The overall prevalence rates for each year will serve as the time-dependent measurement variable of the growth mixture models discussed in the following sections.

**Table 2:** Descriptive statistics of the overall prevalence rates (Prev) for four panel waves.

| Index          | $\bar{x}$ | $s$  | % Zeros |
|----------------|-----------|------|---------|
| Prev ( $t_1$ ) | 0.46      | 1.16 | 76.9    |
| Prev ( $t_2$ ) | 0.69      | 1.42 | 68.3    |
| Prev ( $t_3$ ) | 0.86      | 1.65 | 63.5    |
| Prev ( $t_4$ ) | 0.82      | 1.59 | 62.9    |

**Table 3:** Correlation matrix of the overall prevalence rates and the exogenous variables.

|                | Prev ( $t_1$ ) | Prev ( $t_2$ ) | Prev ( $t_3$ ) | Prev ( $t_4$ ) | Gender | Educ. Level |
|----------------|----------------|----------------|----------------|----------------|--------|-------------|
| Prev ( $t_1$ ) | 1.000          |                |                |                |        |             |
| Prev ( $t_2$ ) | 0.542          | 1.000          |                |                |        |             |
| Prev ( $t_3$ ) | 0.406          | 0.629          | 1.000          |                |        |             |
| Prev ( $t_4$ ) | 0.346          | 0.518          | 0.606          | 1.000          |        |             |
| Gender         | -0.131         | -0.146         | -0.189         | -0.171         | 1.000  |             |
| Educ. Level    | -0.110         | -0.173         | -0.138         | -0.171         | -0.002 | 1.000       |

## 4 Latent class growth analysis

The descriptive results suggest that occurrence of delinquency is related to heterogeneity which can be detected with the observed variables under study. Scores of the overall prevalence rates could be used to assign individuals into a group taxonomy similar to the study of Haapasalo and Tremblay (1994). They defined rules based on the frequencies and trends of the subjects' aggression level. Such assignment rules might be reasonable but the individual classification to the various groups is deterministic. The uncertainty of the group membership cannot be quantified in the form of probabilities. Furthermore, the existence of distinct developmental trajectories must be assumed a priori without any test and the chance to create groups with random variation in the trajectories cannot be excluded. Reanalyses of the study of Haapasalo und Tremblay (1994) with latent class growth analysis (LCGA) provided statistical support for only three of the original four groups obtained by deterministic assignment rules (Nagin and Tremblay, 1999). The advantage of the probabilistic classification is obvious.

The exploration of different growth trajectories with the aforementioned panel data (Section 3) will use the advantages of the mixture models. In this section the latent class growth model serve as a preliminary step exploring different developmental processes of deviant behavior and delinquency: First, unconditional models with  $K$  classes using the Poisson and the zero-inflated Poisson distribution of the overall prevalence rates are calculated to detect differences in terms of model fit. In addition, the necessity to include a quadratic growth term into the models is proved (Section 4.1). Second, conditional models incorporating background covariates gender and educational level are tested. These analyses will focus on the observed heterogeneity of the data, i. e., a test of substantive

**Table 4:** Fit of the unconditional LCGA-models with different classes.

| LCGA-Poisson    |                    |      |      |      |      |      |
|-----------------|--------------------|------|------|------|------|------|
| Type            | Test               | C1   | C2   | C3   | C4   | C5   |
| Linear model    | BIC                | 8994 | 6853 | 6540 | 6523 | 6505 |
|                 | Adj. BIC           | 8987 | 6837 | 6514 | 6488 | 6461 |
|                 | LMR-LRT<br>p-value | –    | 2059 | 317  | 35   | 32   |
| Quadratic model | BIC                | 8970 | 6835 | 6528 | 6510 | 6477 |
|                 | Adj. BIC           | 8961 | 6813 | 6493 | 6463 | 6417 |
|                 | LMR-LRT<br>p-value | –    | 2084 | 322  | 43   | 58   |
| LCGA-ZIP        |                    |      |      |      |      |      |
| Type            | Test               | C1   | C2   | C3   | C4   | C5   |
| Linear model    | BIC                | 7613 | 6724 | 6484 | 6473 | 6477 |
|                 | Adj. BIC           | 7601 | 6702 | 6452 | 6431 | 6426 |
|                 | LMR-LRT<br>p-value | –    | 866  | 248  | 30   | 15   |
| Quadratic model | BIC                | 7613 | 6734 | 6500 | 6494 | 6489 |
|                 | Adj. BIC           | 7594 | 6702 | 6456 | 6437 | 6420 |
|                 | LMR-LRT<br>p-value | –    | 873  | 251  | 32   | 29   |

relationships between the covariates and the latent classes as well as the growth variables (Section 4.2).

#### 4.1 Unconditional models

Latent class growth analysis of the unconditional model is performed up to five classes as a linear and as a quadratic growth curve model as well. A quadratic model assumes that the development is not linear, i. e., deviance and delinquency would decrease for certain groups of offenders after a period of an increase. This process was detected in most of the criminological panel studies and can also be expected with our data.

Assuming a Poisson distribution (abbreviated LCGA-Poisson) the BIC and the adjusted BIC favor a four- or a five-class model whereas the LMR-LRT points to a three-class model (cf. first part of Table 4). The LMR-LRT p-values of the four- and five-class model indicate that the addition of classes is not a significant improvement. Assuming a zero-inflated Poisson distribution (abbreviated LCGA-ZIP) the BIC and the adjusted BIC favor again a four- or a five-class model. But again, the LMR-LRT p-values indicate that a fourth and fifth class is redundant (cf. second part of Table 4). The correction of the zero-inflation gives a better representation of the data in all model variants. Therefore, further analyses will focus on the mixture models assuming the zero-inflated Poisson distribution of the dependent variable.

**Table 5:** Distributions of the classes.

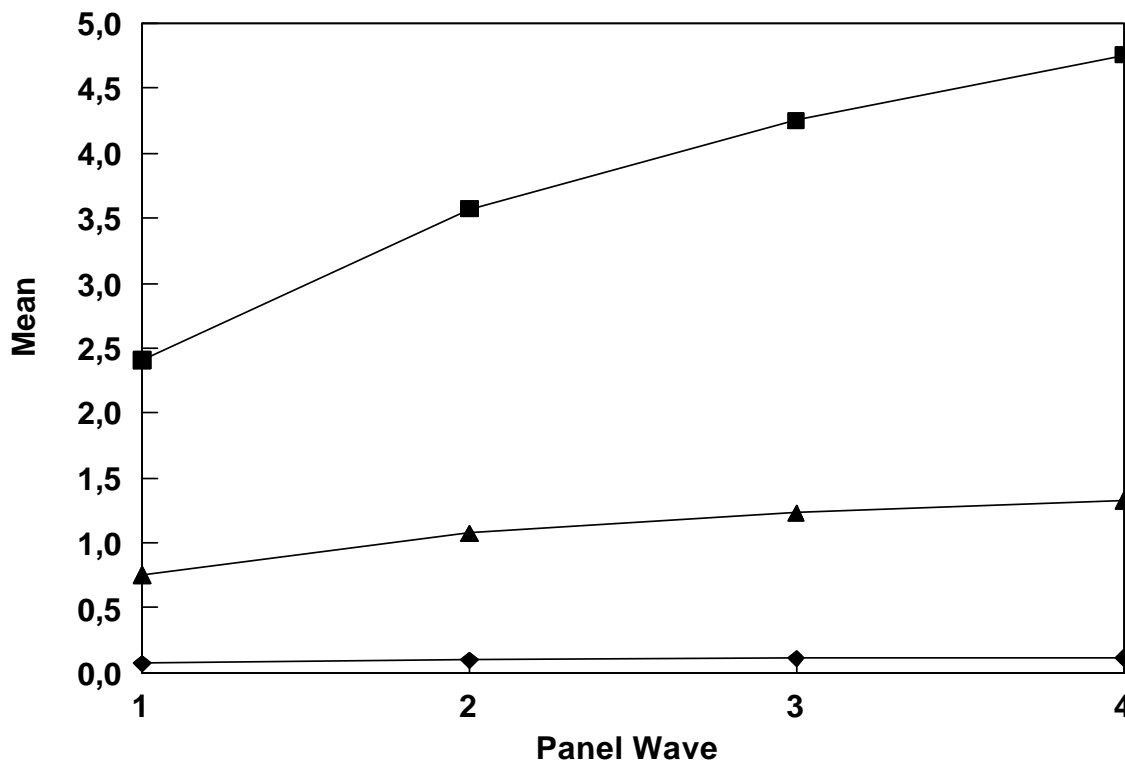
| Classes | ZIP |        | Quadrat-ZIP |        |
|---------|-----|--------|-------------|--------|
|         | n   | (%)    | n           | (%)    |
| 1       | 262 | (32.3) | 265         | (32.6) |
| 2       | 63  | (7.7)  | 64          | (7.9)  |
| 3       | 488 | (60.0) | 484         | (59.5) |

Table 5 shows the class distribution for the zero-inflated (ZIP) and the quadratic zero-inflated model (Quadrat-ZIP) with three classes. The criterion to be in a given class is the most likely latent class membership. For both models the distributions are nearly equal. About 32% of the respondents belong to the first, 8% to the second and 60% to the third class. If both class memberships are crosstabulated, only seven persons are in different classes. According to the entropy measure  $E_k$  (cf. Equation 2.5) the linear and the quadratic ZIP-model indicate a reasonable classification (in both models  $E_k = 0.81$ ).

The estimated growth trajectories of the LCGA-ZIP-model are shown in Figure 3. The lines next to the x-axes represents the non-offenders with almost no deviant and delinquent behavior over the observed time period. This class is the largest one (Class 3 in Table 5). The second largest class, class 1, shows a small development of offending starting from a low level. The estimated means range from 0.75 ( $t_1$ ) to 1.33 ( $t_4$ ). According to the mixture analyses of D'Unger et al. (1998) this class can be characterized as low-rate adolescents (intercept=0.124; slope=0.055). The smallest class, class 2, reflects a somewhat stronger development of deviant and delinquent behavior starting from a significant higher level in the first wave (intercept=1.298; slope=0.092). This class can be characterized as high-rate adolescents. The estimated means range from 2.41 ( $t_1$ ) to 4.75 ( $t_4$ ).

Following Kreuter (2004) the intercept for the inflation part of the model is fixed to zero in all classes. The estimate of the slope is restricted to be equal across classes. In addition, intercepts of the manifest variables ( $y_1^i$  to  $y_4^i$ , cf. Figure 2) are set to be equal within and across classes. The estimated inflation probability decreases from 0.34 (wave 1) to less than 0.10 (wave 4). Differences between the linear and the quadratic LCGA-ZIP-model are small because non-linearity in growth is only observed after the third wave which is due to a slight decrease of offending in classes 1 and 2. With more panel waves the quadratic model specification would give a better representation of the data (cf. the three-class LCGA for the Cambridge study in Muthén, 2004).

Results of the categorical part of the linear LCGA-ZIP-model are discussed as follows (cf. Equation 2.3). Without any exogenous variables, the exponential function of the logit parameter (intercept) is simply the odds ratio for being in the particular class versus the reference class. The class of non-offenders will serve as the reference class. For the first class (low-rate adolescents) a logit coefficient of -0.649 is estimated. The exponentiated value is 0.522. For the second class (high-rate adolescents) the estimated logit coefficient is -2.026 and the exponentiated value is 0.132. Following Equation 2.4 the latent class probability is 0.32 for the first, 0.08 for the second and  $1 - 0.32 - 0.08 = 0.60$  for the third (reference) class. The distribution of those probabilities are equivalent to the distribution of classes given in Table 5.



**Figure 3:** Three-class LCGA-ZIP for four panel waves.

**Table 6:** Fit of the conditional LCGA-ZIP-models with different classes.

| Type            | Test     | C1   | C2   | C3   | C4   | C5   |
|-----------------|----------|------|------|------|------|------|
| Linear Model    | BIC      | 7451 | 6665 | 6444 | 6453 | 6468 |
|                 | Adj. BIC | 7425 | 6624 | 6387 | 6380 | 6379 |
|                 | LMR-LRT  | –    | 796  | 247  | 23   | 19   |
|                 | P-Value  | –    | 0.00 | 0.00 | 0.49 | 0.68 |
| Quadratic Model | BIC      | 7463 | 6688 | 6474 | 6470 | 6471 |
|                 | Adj. BIC | 7425 | 6630 | 6398 | 6375 | 6357 |
|                 | LMR-LRT  | –    | 796  | 248  | 43   | 38   |
|                 | P-Value  | –    | 0.00 | 0.00 | 0.05 | 0.06 |

## 4.2 Conditional models

The aforementioned analyses of the unconditional latent class growth curve model are the basis for the conditional models. Therefore, gender and level of education are included as exogenous variables for the latent class distribution. The fit of the linear and quadratic specification of the latent class growth curve model is again compared up to five classes (cf. Table 6). Similar to the unconditional models the linear specification results in a better model fit than the quadratic one. The BIC, adjusted BIC and the LMR-LRT point to a three-class solution with linear growth.

The estimated class distribution of the conditional linear LCGA-ZIP-model is compa-

rable to the unconditional model: 58.3% of the respondents belong to the class of non-offenders (Class 3), 32.6% to the low-rate adolescents (Class 2) and 9.1% to the high-rate adolescents (Class 1). If the class membership of the three-class unconditional model is crosstabulated with the conditional model, 14 persons move from class 3 to class 2 and 11 persons from class 2 to class 1. According to the entropy measure  $E_k$ , the classification is slightly worse compared to the unconditional model, but still reasonable ( $E_k = 0.72$ ). The estimated parameters differ only slightly from the unconditional specification. The class of high-rate adolescents (intercept=1.582; slope=0.118) indicates that the larger increase of deviant and delinquent behavior gets along with an higher initial rate compared to the class of low-rate adolescents (intercept=0.409; slope=0.053).

According to Figure 1 and Muthén (2004: 355) the regressions of the growth curve variables on the covariates should be allowed unless there are theoretical reasons not to do it. Here, those regression coefficients are restricted to be equal across classes. A substantive effect is only observed for the regression of the intercept on educational level (-0.218), which is in accordance with our hypothesis: The higher the level of education, the lower the initial rate of offending.

Table 7 summarizes the logit part of the conditional linear ZIP-model: the logit parameters, the exponentiated values of the logits (odds) and the latent class probabilities. Equations 2.3 and 2.4 are used for calculations.

**Table 7:** Logit, odds ( $e^{\text{logit}}$ ) and latent class probabilities (Prob.) of the conditional linear ZIP-model (3 classes).

| Exogenous Variable                  | Class | Logit  | $e^{\text{logit}}$ | Prob. |
|-------------------------------------|-------|--------|--------------------|-------|
| Male and low educational level      | 1     | -0.769 | 0.463              | 0.159 |
|                                     | 2     | 0.377  | 1.458              | 0.499 |
|                                     | 3     | 0      | 1.0                | 0.342 |
| Male and medium educational level   | 1     | -1.113 | 0.329              | 0.142 |
|                                     | 2     | -0.006 | 0.994              | 0.428 |
|                                     | 3     | 0      | 1.0                | 0.430 |
| Male and high educational level     | 1     | -1.457 | 0.233              | 0.122 |
|                                     | 2     | -0.389 | 0.678              | 0.355 |
|                                     | 3     | 0      | 1.0                | 0.523 |
| Female and low educational level    | 1     | -1.902 | 0.149              | 0.080 |
|                                     | 2     | -0.345 | 0.708              | 0.381 |
|                                     | 3     | 0      | 1.0                | 0.539 |
| Female and medium educational level | 1     | -2.250 | 0.106              | 0.067 |
|                                     | 2     | -0.728 | 0.483              | 0.304 |
|                                     | 3     | 0      | 1.0                | 0.629 |
| Female and high educational level   | 1     | -2.590 | 0.075              | 0.053 |
|                                     | 2     | -1.111 | 0.329              | 0.234 |
|                                     | 3     | 0      | 1.0                | 0.713 |

Gender is coded 0 and 1 (male/female), educational level is coded 0, 1 and 2 (low, medium, high). Class 1 is the high-rated class, class 2 the low-rated, and class 3 the zero class.

The probability for males with low educational level is nearly 16% to be in class 1 (high-rate adolescents), about 50% to be in class 2 (low-rate adolescents) and only about 34% to be in class 3 (non-offenders). The probability to be classified in class 3 increases with the educational level of the school. The probability for females with low educational level is only 8% to be in class 1, nearly 38% to be in class 2, but 54% to be in class 3. As with the males, the probability to be classified in class 3 increases with the educational level. But this increase for the females is much stronger compared to the males. The chance to be a member of the high- or low-rate class is clearly greater for males than females within all levels of education. On the average, the effect of the latent classes on gender is stronger than the effect on educational level.

## 5 Growth mixture models

It has been shown, that unobserved as well as observed heterogeneity can be considered simultaneously with LCGA to detect different trajectories (classes) of deviant and delinquent behavior. Time-invariant variables, like gender and level of education, are able to predict the probabilities to belong to specific classes. Similar analyses are reported with other longitudinal studies on criminal offending (e. g. Nagin and Land, 1993; Nagin, 1999). But LCGA is only a submodel of the growth mixture model (GMM) that allows the variability of the growth parameters within given classes. The model analyses of the previous section has shown that linear models incorporating the zero-inflated Poisson distribution have the best model fits. Therefore, the zero-inflated Poisson distribution will be used for estimation of the growth mixture models (abbreviated GMM-ZIP). Unconditional models are discussed in Section 5.1 followed by the conditional ones including the exogenous variables gender and educational level (Section 5.2).

### 5.1 Unconditional models

In a step-by-step procedure the unconditional GMM-ZIP models include either the variation of the intercepts within the classes or the variation of the slopes within the classes (with the restriction of equal intercept or slope variances across classes). Both model variants have equal degrees of freedom and fit measures can be compared easily. Table 8 provides the model results up to four classes. In general, models with random intercepts have a slightly better model fit than models with random slopes. According to the BIC and the LMR-LRT a three-class model is sufficient for the data. All four-class model are rejected by the LMR-LRT. Results in detail show, that in each solution one class contains less than six persons.

If the intercept variance of the class of non-offenders is set to zero (cf. a similar specification in Muthén, 2004: 362), the model fit gets worse compared to the previous specification (the three-class specification is rejected with LMR-LRT=3 and  $p=0.65$ ; see first model C3a in Table 8).<sup>5</sup> Alternatively, if the slope variance of the class of non-offenders is set to zero, the model fit is nearly equal compared to the corresponding three-class

<sup>5</sup>Alternatively, means and variances of the growth curve variables of the non-offender class can be fixed to zero. In the same class the means of the manifest variables are fixed to the value  $-15.00$  (cf. the specification in Kreuter, 2004). Again, the model fit decreases similar to model C3a.



**Table 8:** Fit of the unconditional growth mixture models (GMM-ZIP) with different classes.

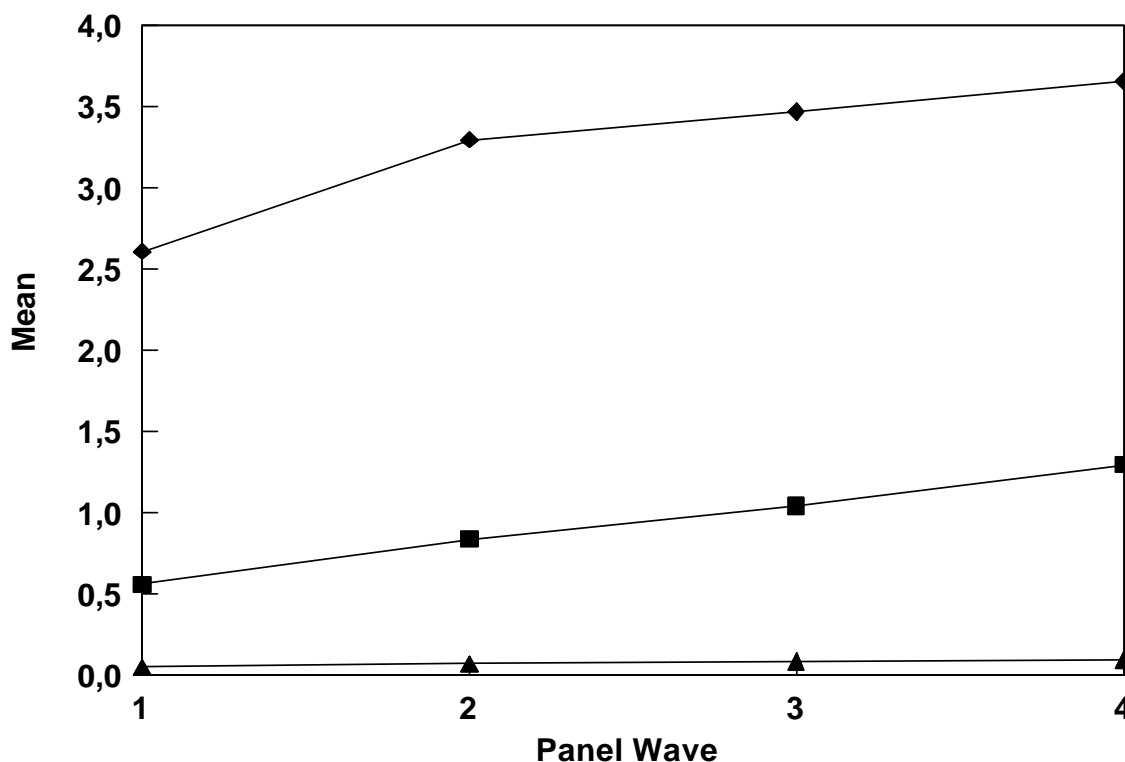
| Type             | Test               | C1   | C2          | C3         | C3a        | C4         |
|------------------|--------------------|------|-------------|------------|------------|------------|
| Random intercept | BIC                | 6464 | 6450        | 6431       | 6461       | 6445       |
|                  | Adj. BIC           | 6448 | 6424        | 6396       | 6426       | 6401       |
|                  | LMR-LRT<br>p-value | –    | 33<br>0.03  | 37<br>0.00 | 3<br>0.65  | 5<br>0.17  |
| Random slope     | BIC                | 6872 | 6522        | 6442       | 6444       | 6447       |
|                  | Adj. BIC           | 6856 | 6497        | 6408       | 6409       | 6402       |
|                  | LMR-LRT<br>p-value | –    | 352<br>0.00 | 95<br>0.00 | 94<br>0.00 | 15<br>0.62 |

Model C3a includes the restriction that the intercept or the slope of the zero class has a variance of zero.

solution (the three-class specification is confirmed with LMR-LRT=94 and  $p=0.00$ ; see second model C3a in Table 8). Due to the fact that this model specification is more parsimonious than the previous one, it will be accepted for further substantive interpretations. Though, a random slope is estimated for the second and third class (low- and high-rated adolescents) while the variance of the slope in the first class (non-offenders) are fixed to zero. It should also be noted, that the entropy measure is much better in all random slope models (average  $E_K = 0.70$ ) compared to the random intercept models. One more general specification is the growth mixture model with random intercept and random slope estimated simultaneously. But, even with several alternative starting values a solution of this model with three classes could not be obtained.

Due to the most likely latent class membership 58% of the respondents are classified in the first ( $N=473$ ), 33% into the second ( $N=270$ ) and 9% into the third class ( $N=70$ ). The sequence fits into the order of the class of non-offenders, the class of low-rate adolescents and finally, the class of high-rate adolescents. If the class membership is crosstabulated with the unconditional LCGA-model discussed in Section 4.1, most of the cases remain in the particular classes. Only 16 persons move from the low-rate class to the high-rate class, 9 persons move vice versa. Another 15 persons move from the non-offenders class to the low-rate class. All in all, the zero class is smaller in the GMM-ZIP-model than in the corresponding LCGA-ZIP-model whereas the other classes have higher proportions in the GMM-ZIP-model. This result is comparable to the mixture analyses of the Cambridge data in Muthén (2004: 362).

The estimated growth trajectories of the GMM-ZIP model with random slopes are shown in Figure 4. The line next to the x-axis represents the non-offenders (Class 1) with almost no deviant and delinquent behavior over the observed time period. This class is again the largest class. The second largest class 2 shows a small development of offending starting from a low level (intercept=-0.229; slope=0.088). The estimated means range from 0.54 (wave 1) to 1.29 (wave 4). The smallest class 3 reflects a somewhat stronger development starting from a higher level in the first wave (intercept=1.316; slope=0.080). The estimated means range from 2.61 (wave 1) to 3.66 (wave 4). The three growth trajectories of the GMM-ZIP model are similar to those of the LCGA-ZIP model (Figure 3), the estimated slope variance restricted to be equal across classes is significant (vari-



**Figure 4:** Three-class GMM-ZIP for four panel waves.

ance=0.049;  $z=16.27$ ).

Results of the categorical part of the unconditional GMM-ZIP model are also comparable to the unconditional LCGA-ZIP model. In the GMM-ZIP model the class of high-rate adolescents will serve as the reference class. For the first class (non-offenders) a logit coefficient of 1.656 is estimated. For the second class (low-rate adolescents) the estimated logit coefficient is 1.256. Following Equation 2.4 the latent class probability is 0.54 for the first, 0.36 for the second and 0.10 for the third (reference) class.

## 5.2 Conditional models

The unconditional GMM-ZIP model can now be extended by the exogenous variables gender and level of education. In a step-by-step procedure the conditional GMM-ZIP models include either the variation of the intercepts within the classes or the variation of the slopes within the classes (with the restriction of equal intercept or slope variances across classes). Table 9 provides the model results up to four classes. According to the BIC and the LMR-LRT, three-class models with random intercepts or random slopes are sufficient for the data. Both four-class model are rejected by the LMR-LRT.

If the intercept variance of the class of non-offenders is set to zero, the model fit gets worse according to the BIC. But different from the unconditional specification, the three-class solution is not rejected by the LMR-LRT (LMR-LRT=41 with  $p=0.00$ ; see first model C3a in Table 9). Alternatively, if the slope variance of the class of non-offenders is set to zero, the model fit is nearly equal compared to the unrestricted three-class so-

**Table 9:** Fit of the conditional GMM-ZIP-models with different classes.

| Type             | Test               | C1   | C2          | C3         | C3a         | C4         |
|------------------|--------------------|------|-------------|------------|-------------|------------|
| Random intercept | BIC                | 6402 | 6400        | 6395       | 6414        | 6421       |
|                  | Adj. BIC           | 6380 | 6362        | 6341       | 6361        | 6351       |
|                  | LMR-LRT<br>p-value | –    | 34<br>0.02  | 32<br>0.06 | 41<br>0.00  | 8<br>0.43  |
| Random slope     | BIC                | 6816 | 6469        | 6406       | 6408        | 6413       |
|                  | Adj. BIC           | 6794 | 6431        | 6352       | 6354        | 6353       |
|                  | LMR-LRT<br>p-value | –    | 369<br>0.00 | 94<br>0.00 | 206<br>0.00 | 16<br>0.66 |

Model C3a includes the restriction that the intercept or the slope of the zero class has a variance of zero.

lution (the BIC value difference between the models is only 2). The restricted model is also confirmed by the LMR-LRT (LMR-LRT=206 with  $p=0.00$ ; see second model C3a in Table 9). So, this model specification will be accepted for further substantive interpretations. Equal to the restricted unconditional GMM-ZIP model (cf. Table 8), a random slope is estimated for the second and third class (low- and high-rated adolescents) while the variance of the slope in the first class (non-offenders) is fixed to zero. According to the most likely class membership 58.2% of the respondents belong to the class of non-offenders (Class 1), 33.2% to the low-rate adolescents (Class 2) and 8.5% to the high-rate adolescents (Class 3).

The estimated class distribution of the conditional linear GMM-ZIP model can be compared to the conditional linear LCGA-ZIP model in Section 4.2 via crosstabulation of both membership files: 16 persons move from the low to the high-rated class, 9 persons vice versa and 15 persons from the non-offender class to the low-rated class. All other non-offenders of the GMM-ZIP model are also non-offenders in LCGA-ZIP model. The entropy measure of this model is reasonable ( $E_K = 0.71$ ) and nearly the same as for the conditional LCGA-ZIP model. The regression coefficients of the growth curve variables on the exogenous variables are again restricted to be equal across classes. No substantive effects are observed.

The categorical part of the conditional GMM-ZIP model is summarized in Table 10. The probability for males with low educational level to be in class 3 (high-rate adolescents) is 29%, to be in class 2 (low-rated adolescents) is about 46% and to be in class 1 (non-offenders) is only 25%. The probability to be classified as non-offenders increases with the educational level of the school. The probability for females with low educational level is about 12% to be in class 3, nearly 42% to be in class 2 and 46% to be in class 3. As with the males, the probability to be classified as non-offenders increases with the educational level. But, compared to the males the increase is much stronger. The chance to be a member of the high- or low-rate class is clearly greater for males than females within all levels of education. Comparable to the conditional LCGA-ZIP model (cf. Table 7), the effect of gender on the latent classes is stronger than the effect on educational level. One major difference between the conditional LCGA-ZIP model and the conditional GMM-ZIP model is the distribution of the latent class probabilities within a com-

**Table 10:** Logit, odds ( $e^{\text{logit}}$ ) and latent class probabilities (Prob.) of the conditional linear ZIP-model (3 classes).

| Exogenous Variable                  | Class | Logit  | $e^{\text{logit}}$ | Prob. |
|-------------------------------------|-------|--------|--------------------|-------|
| Male and low educational level      | 1     | -0.159 | 0.853              | 0.247 |
|                                     | 2     | 0.467  | 1.595              | 0.463 |
|                                     | 3     | 0      | 1.0                | 0.290 |
| Male and medium educational level   | 1     | 0.700  | 2.014              | 0.372 |
|                                     | 2     | 0.876  | 2.401              | 0.443 |
|                                     | 3     | 0      | 1.0                | 0.185 |
| Male and high educational level     | 1     | 1.559  | 4.754              | 0.502 |
|                                     | 2     | 1.285  | 3.615              | 0.386 |
|                                     | 3     | 0      | 1.0                | 0.107 |
| Female and low educational level    | 1     | 1.360  | 3.896              | 0.464 |
|                                     | 2     | 1.254  | 3.504              | 0.417 |
|                                     | 3     | 0      | 1.0                | 0.119 |
| Female and medium educational level | 1     | 2.219  | 9.198              | 0.594 |
|                                     | 2     | 1.663  | 5.275              | 0.341 |
|                                     | 3     | 0      | 1.0                | 0.065 |
| Female and high educational level   | 1     | 3.078  | 21.714             | 0.708 |
|                                     | 2     | 2.072  | 7.941              | 0.259 |
|                                     | 3     | 0      | 1.0                | 0.033 |

Gender is coded 0 and 1 (male/female), educational level is coded 0, 1 and 2 (low, medium, high). Class 1 is the zero class, class 2 the low-rated, and class 3 the high-rated class.

combination of exogenous variables' categories. For example, the probability to belong to the high-rate adolescents is for males higher in the GMM-ZIP model than in the LCGA-ZIP model. The significant variance estimate of the slope (variance=0.048;  $z=15.71$ ) leads to a more adequate classification of the respondents in the GMM-ZIP model and supports the assumption of individual growth variation within the classes.

## 6 Discussion

The general framework of growth mixture modeling outlined by Muthén (2002, 2004) integrates several approaches of longitudinal growth modeling, e. g., the semiparametric group-based model developed and applied by Nagin and Land (1993) Nagin (1999), and Jones et al. (2001). This model is equivalent to a latent class growth model which fixes the intercept and slope variances to zero. According to an easier estimation of the parameters the latent class growth analysis is useful for a first evaluation of the unobserved heterogeneity in the data. After these analyses the variability of the class specific intercepts and slopes can be studied with the more general growth mixture model. If count data with largely skewed distributions to zero are analyzed, the assumption of continuously distributed variables can be replaced by the Poisson or the zero-inflated Poisson

model. *Mplus* (Version 3.11) allows those tests of growth mixture models assuming different distributions of the manifest variables under study.

Data from a four-wave panel study of adolescents are used to study unobserved heterogeneity in the development of deviant and delinquent behavior. Starting with an unconditional latent class growth model three classes can be obtained: Non-offenders, low-rate adolescents and high-rate adolescents. Considering the zero-inflation of the measurements results in a better model fit. Due to the limited number of panel waves in our data (4 waves) a linear growth specification is sufficient. This result is also confirmed by a growth mixture specification allowing random slopes in the low- and high-rated classes. The conditional specifications of the mixture models include gender and educational level of the schools. These exogenous variables are related to the latent class distribution: Male adolescents with an low educational level on schools are more likely to be in the high-rated class than female adolescents on the same educational level. And female adolescents are more likely to be non-offenders than male adolescents. These results confirm our hypotheses and they are similar compared to other criminological studies using mixed Poisson models with panel data of adolescents' deviant and delinquent behavior (D'Unger et al., 1998; Jones et al., 2001). Their models, equivalent to LCGA, assume zero variances of the intercept and slope variable. Therefore, our analyses with the growth mixture model are more general even if it was only possible to estimate intercept and slope variances separately. The analyses do not confirm a convex trajectory that declines from a high initial rate which was found by McDermott and Nagin (2001) using the National Youth Survey. A similar trajectory can only be obtained in a five-class solution which is rejected by several model fit criterias (BIC, adjusted BIC and LMR-LRT). All in all, the stepwise procedure from latent growth analysis to growth mixture modeling has shown the capabilities of longitudinal developmental analyses where individual growth trajectories are heterogenous and belong to a finite number of unobserved populations.

The present analyses are limited due to the four-wave panel design of our study. The mixture models were only conditioned on time-independent predictors, gender and educational level. Note, that educational level can also be analyzed as a time-dependent exogenous variable regarding a certain mobility of the adolescents between different schools. Actually, only a minor proportion of adolescents changed between the schools. In Germany, the proportion of school changes increases significantly after the 10th grade, which is beyond the period of our data collection. Several studies have considered various predictors in their studies, like socioeconomic status of the parents, parental behavior, moral beliefs, attitudes about crime or delinquent peers (Nagin, 1999; McDermott and Nagin, 2001). Therefore, further growth mixture model explorations should also include substantively relevant time-dependent variables to test parallel developments and their interrelationships over time. These models have to cope with different assumptions about the measurement of the variables (continuous and count data) which are easy to implement in *Mplus*. Actually, those extensions are beyond the scope of the present article.

## Acknowledgments

The author would like to thank the 2005 Applied Statistics conference organizers for the opportunity to present parts of the research work on mixture models. The author

would also like to thank Dr. Herbert Matschinger for helpful comments about specification details in the program *Mplus*.

## References

- [1] Bauer, D.J. and Curran, P.J. (2003): Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, **8**, 338-363.
- [2] Bauer, D.J. and Curran, P.J. (2004): The integration of continuous and discrete and latent variable models: Potential problems and possible solutions. *Psychological Methods*, **9**, 3-29.
- [3] Akaike, H. (1987): Factor analysis and the AIC. *Psychometrika*, **52**, 317-332.
- [4] Bentler, P.M. (2001): *EQS 6: Structural Equations Program Manual*. Encino: Multivariate Software.
- [5] Boers, K., Reinecke, J., Motzke, K., and Wittenberg, J. (2002): Wertorientierungen, Freizeitstile und Jugenddelinquenz. *Neue Kriminalpolitik*, **14**, 141-146.
- [6] Curran, P.J. and Hussong, A.M. (2002): Modeling intraindividual variability with repeated measures data: Method and applications. In D.S. Moskowitz and S.L. Hershberger (Eds.): *Structural equation modeling of repeated measures data: Latent curve analysis*, 59-85. Mahwah: Lawrence Erlbaum.
- [7] D'Unger, A., Land, K.C., McCall, P.L., and Nagin, D.S. (1998): How many latent classes of delinquent/criminal careers? Results from mixed Poisson regression analyses of the London, Philadelphia and Racine cohort studies. *American Journal of Sociology*, **103**, 1593-1630.
- [8] Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977): Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.
- [9] Duncan, T.E., Duncan, S.C., Strycker, L.A., Li, F., and Alpert, A. (1999): *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah: Lawrence Erlbaum.
- [10] Farrington, D.P. and West, D.J. (1990): The Cambridge study in delinquent development: A longterm follow-up of 411 London males. In H.J. Kerner and G. Kaiser (Eds.): *Kriminalität: Persönlichkeit, Lebensgeschichte und Verhalten*, 115-138. Berlin: Springer.
- [11] Haapasalo, J. and Tremblay, R.E. (1994): Physically aggressive boys from ages 6 to 12: Family background, parenting behavior, and prediction of delinquency. *Journal of Consulting and Clinical Psychology*, **62**, 1044-1052.

- 
- [12] Hox, J.J. (2002): *Multilevel analysis. Techniques and applications*. Mahwah: Lawrence Erlbaum.
- [13] Jeffries, N.O. (2003): A note on 'testing the number of components in a normal mixture'. *Biometrika*, **90**, 991-994.
- [14] Jones, B.L., Nagin, D.S., and Roeder, K. (2001): A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods & Research*, **29**, 374-393.
- [15] Jöreskog, K.G. and Sörbom, D. (2004): *LISREL 8.7 for Windows*. Lincolnwood: Scientific Software International, Inc.
- [16] Kreuter, F. (2004): *Longitudinal models with zero inflation*. Unpublished paper. University of California at Los Angeles, Graduate School of Education and Information Studies.
- [17] Lambert, D. (1992): Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, **34**, 1-13.
- [18] Land, K.C., McCall, P.L., and Nagin, D.S. (1996): A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models with empirical applications to criminal careers data. *Sociological Methods and Research*, **24**, 387-442.
- [19] Liang, K.-Y. and Zeger, S.L. (1986): Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13-22.
- [20] Lo, Y., Mendell, N.R., and Rubin, D.B. (2001): Testing the number of components in a normal mixture. *Biometrika*, **88**, 767-778.
- [21] McArdle, J.J. (1988): Dynamic but structural equation modeling of repeated measures data. In J.R. Nesselroade and R.B. Cattell (Eds.): *Handbook of multivariate experimental psychology*, 561-614. New York: Plenum.
- [22] McArdle, J.J. and Epstein, D. (1987): Latent growth curves within developmental structural equation models. *Child Development*, **58**, 110-133.
- [23] McDermott, S. and Nagin, D. S. (2001): Same or different? Comparing offender groups and covariates over time. *Sociological Methods and Research*, **29**, 282-318.
- [24] Meredith, M. and Tisak, J. (1990): Latent curve analysis. *Psychometrika*, **55**, 107-122.
- [25] Muthén, B. (1991): Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins and J. Horn (Eds.): *Best Methods for the Analysis of Change*, 1-17. Washington DC: American Psychological Association.
- [26] Muthén, B. (1997): Latent variable modeling with longitudinal and multilevel data. In A. Raftery (ed.): *Sociological Methodology*, 453-480. Boston: Blackwell Publishers.

- [27] Muthén, B. (2001a): Latent variable mixture modeling. In G.A. Marcoulides and R.E. Schumacker (Eds.): *New Developments and Techniques in Structural Equation Modeling*, 1-33. Mahwah: Lawrence Erlbaum.
- [28] Muthén, B. (2001b): Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. and Sayer, A. (Eds.): *New Methods for the Analysis of Change*, 291-322. Washington, D.C.: APA.
- [29] Muthén, B.O. (2002): Beyond SEM: General latent variable modeling. *Behaviorometrika*, **29**, 81-117.
- [30] Muthén, B.O. (2004): Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.): *The Sage Handbook of Quantitative Methodology for the Social Sciences*, 345-368. Thousand Oaks: Sage.
- [31] Muthén, B.O and Curran, P. J. (1997): General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, **2**, 371-402.
- [32] Muthén, B. and Shedden, K. (1999): Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, **55**, 463-469.
- [33] Muthén, L. and Muthén, B.O. (2001): *Mplus: The Comprehensive Modeling Program for Applied Researchers: User's guide* (2nd ed.). Los Angeles: Muthén and Muthén.
- [34] Muthén, L. and Muthén, B.O. (2004): *Mplus User's Guide* (3rd ed.). Los Angeles: Muthén and Muthén.
- [35] Nagin, D.S. (1999): Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods*, **4**, 139-157.
- [36] Nagin, D.S. (2005): *Group-based Modeling of Development*. Cambridge: Harvard University Press.
- [37] Nagin, D.S. and Land, K.C. (1993): Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology*, **31**, 327-362.
- [38] Nagin, D.S. and Tremblay, R.E. (1999): Trajectories of boys physical aggression, opposition and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, **70**, 1181-1196.
- [39] Pöge, A. (2005): Persönliche Codes bei Längsschnittstudien: Ein Erfahrungsbericht. *ZA-Information*, **8**, 50-69.
- [40] Rao, C.R. (1958): Some statistical methods for comparison of growth curves. *Biometrics*, **14**, 1-17.



- 
- [41] Roeder, K., Lynch, K.G., and Nagin, D.S. (1999): Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association*, **94**, 766-776.
- [42] Ross, S.M. (1993): *Introduction to Probability Models* (5th ed.). New York: Academic Press.
- [43] Rudas, T., Clogg, C.C., and Lindsay, B.G. (1994): A new index of fit based on mixture methods for the analysis of contingency tables. *Journal of the Royal Statistical Society, Series B*, **56**, 623-639.
- [44] Schwarz, G. (1978): Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.
- [45] Tracy, P.E., Wolfgang, M.E., and Figlio, R.M. (1990): *Delinquency in Two Birth Cohorts*. New York: Plenum.
- [46] Tremblay, R.E., Desmarais-Gervais, L., Gagnon, C., and Charlebois, P. (1987): The preschool behavior questionnaire: Stability of its factor structure between cultures, sexes, ages and socioeconomic classes. *International Journal of Behavioral Development*, **10**, 467-484.
- [47] Tucker, L.R. (1958): Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19-23.
- [48] Yang, C.C. (1998): *Finite mixture model selection with psychometric applications*. Unpublished doctoral dissertation, University of Groningen.