Stage-Sequential Growth Mixture Modeling of Criminological Panel Data

Jost Reinecke, Maike Meyer, and Klaus Boers

Abstract The detection of distinctive developmental trajectories is of great importance in criminological research. The methodology of growth curve and finite mixture modeling provides the opportunity to examine different developments of offending. With latent growth curve models (LGM) (Meredith and Tisak, Psychometrika 55:107–122, 1990) the structural equation methodology offers a strategy to examine intra- and interindividual developmental processes of delinquent behavior. There might, however, not be a single but a mixture of populations underlying the growth curves which refers to unobserved heterogeneity in the longitudinal data. Growth mixture models (GMM) introduced by Muthén and Shedden (Biometrics 55:463–469, 1999) can consider unobserved heterogeneity when estimating growth curves. GMM distinguish between continuous variables which represent the growth curve model and categorical variables which refer to subgroups that have a common development in the growth process. The models are usually based on single-phase data which associate any event with a specific period. Panel data, however, often contain several relevant phases. In this context, stage-sequential growth mixture models with multiphase longitudinal data become increasingly important. Kim and Kim (Structural Equation Modeling: A Multidisciplinary Journal 19:293–319, 2012) investigated and discussed three distinctive types of stage-sequential growth mixture models: traditional piecewise GMM, discontinuous piecewise GMM, and sequential process GMM. These models will be applied here to examine different stages of delinquent trajectories within the time range of adolescence and young adulthood using data from the German panel study *Crime in the Modern City* (CrimoC, Boers et al., Monatsschrift für Kriminologie und Strafrechtsreform 3:183–202, 2014). Methodological and sub-

J. Reinecke (✉)
Faculty of Sociology, Bielefeld University, Postbox 100131, D-33501 Bielefeld, Germany
e-mail: jost.reinecke@uni-bielefeld.de

M. Meyer
Institute of Political Science, University of Münster, Scharnhorststraße 100, 48151 Münster, Germany

K. Boers
Faculty of Law, Universitätsstr. 14-16, 48143 Münster, Germany

stantive differences between single-phase and multi-phase models are discussed as well as recommendations for future applications.

Introduction

Investigations regarding the development of delinquency during the life course are currently of great importance in longitudinal criminological research. Over the past 20 years a variety of criminologists have argued that there are distinctive groups of offenders which can be described by different delinquent trajectories (Loeber & LeBlanc 1990; Moffitt 1993; Sampson & Laub 2003; Thornberry 2005).

> A trajectory is a pathway or line of development over the life span such as worklife, parenthood, and criminal behavior. Trajectories refer to long-term patterns of behavior and are marked by a sequence of transitions. Transitions are marked by life events (e.g. first job or first marriage) that are embedded in trajectories and evolve over shorter time spans. (Sampson & Laub 1997, p. 142)

Major methodological developments in criminological longitudinal research are influenced by the debate whether distinctive groups of criminal behavior can be identified and in which way the development of a "criminal career" can be incorporated in a statistical model. The debate is mainly enforced by Moffitt's dual taxonomy of offending behavior: The adolescent limited offenders exhibit antisocial behavior only during adolescence while life-course-persistent offenders begin to behave antisocially early in childhood and continue this behavior into adulthood (Moffitt 1993). In further analyses of data from the "Dunedin Multidisciplinary Health and Development Study" (Moffitt, Caspi, Rutter, & Silva 2011) four antisocial behavior trajectory groups were identified among females and males: life-course-persistent, adolescent-onset, childhood-limited, and low trajectory groups (Odgers et al. 2008). Furthermore, Nagin and collaborates explored population heterogeneity in behavioral trajectories using other longitudinal studies, like the "Cambridge Study" (Farrington & West 1990), the "Philadelphia Study" (Tracy, Wolfgang, & Figlio 1990), and the "Montreal Study" (Tremblay, Desmarais-Gervais, Gagnon, & Charlebois 1987). Depending on the type of the dependent variable, nature of the sample, and characteristics of the community, three to five trajectories were detected which reflect different intensity and growth of delinquency. These trajectories distinguish between non-offenders, a time-limited delinquent behavior through adolescence and a more or less chronic group of offenders (D'Unger, Land, McCall, & Nagin 1998; Nagin 1999; Nagin & Land 1993).

The reported findings suggest that there is a variety of heterogeneous trajectories which differ in the age of entry and exit in delinquency, its intensity, its duration, and its continuity (for an overview, see Piquero 2008). Furthermore the research on delinquent trajectories has shown that most people commit delinquent acts rarely or do not become delinquent at all. In most cases early intensive offenders desist from crime. There are, however, trajectories which are marked by high delinquency rates

or by increases in delinquency (Mariotti & Reinecke 2010; Moffitt 1993; Sampson & Laub 2003; Thornberry 2005). Moreover, research has shown that transition points are relevant for the analysis of delinquent trajectories (Sampson & Laub 1993).

Techniques of longitudinal statistical modeling are highly relevant and gained considerable attention in the examination of delinquent trajectories. With latent growth curve models (LGM) (Meredith & Tisak 1990), the structural equation methodology offers a strategy to examine intra- and interindividual developmental processes of delinquent behavior. It can, however, not be assumed that there is always a single population underlying the growth curves. Therefore, observed as well as unobserved heterogeneity has to be taken into account. Observed heterogeneity can be considered by relevant exogenous variables (e.g. gender) which are related to the growth curve variables explaining parts of their variances. To capture unobserved heterogeneity, the latent growth curve model has to be enlarged by a growth mixture model (Muthén & Shedden 1999) which contains the continuous manifest and latent variables as well but in addition categorical variables. The latter ones refer to particular subgroups reflecting different developmental processes. Analyses with a growth mixture model usually assume single-phase data which associate any event with a specific time period. However, longitudinal data often contain transition points which separate different phases of the development under study. An appropriate framework for multi-phase longitudinal data regarding unobserved heterogeneity is the extension of the growth mixture models (GMM) to stage-sequential growth mixture models (Kim & Kim 2012). These models can lead to a better understanding of the developmental process over several phases.

The following section "Method and Models" discusses basic conceptions of growth curve, growth mixture, and the multi-phase growth mixture models as well as the respective methods of model estimation and model evaluation. Additional attention is given to the distributional assumptions of the manifest time-variant variable under study. In the case of count variables Poisson or negative binomial distributions (Hilbe 2011) can be considered which give a better model representation compared to the assumption of a continuous distribution (Reinecke & Seddig 2011).

All models are applied to panel data from the German panel study *Crime in the Modern City* (CrimoC).[1] The data set contains 3938 adolescents and young adults who participated at least twice in a row in the eight panel waves. Data, variables and descriptive statistics are discussed in section "Data, Variables, and Descriptive Statistics".

Results are presented in section "Modeling Results". The analysis starts with single-phase growth mixture models considering up to eight classes and considers various specifications of stage-sequential growth mixture models. Finally, in section "Conclusion" models are compared and discussed with recommendations for further analyses.

---

[1]Principal investigators of the panel study are Klaus Boers (University of Münster) and Jost Reinecke (University of Bielefeld). Since 2002 the study is continuously funded by the German Science Foundation (DFG). Further information can be found at www.crimoc.org.

Method and Models

Latent Growth Curve Models

LGM specified with structural equations have already been discussed in several papers (McArdle 2009; McArdle & Epstein 1987; Meredith & Tisak 1990) and books (Bollen & Curran 2006; Duncan, Duncan, Strycker, Li, & Alpert 2006; Reinecke 2012). As is typical for all structural equation models, growth curve models distinguish between a measurement and a structural model. Structural model refers to the intraindividual development whereas the measurement model refers to interindividual differences of those trends. For a growth curve model with two factors, the measurement model can be formulated as follows:

$$y_t = \lambda_{1t}\eta_1 + \lambda_{2t}\eta_2 + \epsilon_t \tag{1}$$

y_t are the manifest variables at time t, which are related to the latent variables η_1 and η_2. η_1 is the initial level factor or intercept factor while η_2 is the linear growth factor or slope factor. λ_{1t} and λ_{2t} are the factor loadings on η_1 and η_2. ϵ_t is the measurement error of y_t. For each latent variable η, a structural equation has to be formulated as follows:

$$\eta_1 = \alpha_1 + \zeta_1 \tag{2}$$

$$\eta_2 = \alpha_2 + \zeta_2 \tag{3}$$

The latent variables η_1 and η_2 are described by their means (α_1 and α_2) as well as by their residuals (ζ_1 and ζ_2). ζ_1 and ζ_2 can be defined as deviations of the latent variables from their mean values.[2] Variances and covariances of the latent variables are specified in the matrix Ψ:

$$\Psi = \begin{pmatrix} \psi_{11} \\ \psi_{21} \ \psi_{22} \end{pmatrix} \tag{4}$$

Assuming linear growth, the factor loadings for η_1 have to be fixed to one and the factor loadings for η_2 have to be restricted according to a linear development:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 \ 0 \\ 1 \ 1 \\ 1 \ 2 \\ 1 \ 3 \end{bmatrix} * \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} \tag{5}$$

---

[2]The structural equations can be extended by time-invariant latent variables which serve as predictors of the intercept and slope (e.g. gender). Then, ζ_1 and ζ_2 are no longer deviations from the mean values of the latent variables η_1 and η_2 (Reinecke 2012, p. 6).

To model non-linear growth curves it is possible to extend the two-factor model by additional latent variables, for instance a quadratic term. The measurement and structural equations (1)–(3) of the two-factor model described above can be extended as follows:

$$y_t = \lambda_{1t}\eta_1 + \lambda_{2t}\eta_2 + \lambda_{3t}^2\eta_3 + \epsilon_t \tag{6}$$

$$\begin{aligned} \eta_1 &= \alpha_1 + \zeta_1 \\ \eta_2 &= \alpha_2 + \zeta_2 \\ \eta_3 &= \alpha_3 + \zeta_3 \end{aligned} \tag{7}$$

Another possibility to cope with nonlinearity is the so-called piecewise growth curve model which is useful when transition points are assumed across the time range (Bollen & Curran 2006). Such a model contains two or more latent variables. Contrary to the linear growth model, those models can be used to analyze multiphase data. Piecewise growth curve models are meaningful when transition points can be found in the course of development (see, for instance, Raudenbush & Bryk 2002). Assuming one transition point, the first trajectory describes the development between the intercept and the transition point. The second trajectory describes development after the transition point. If six panel waves and a transition point for the third panel wave are assumed, the following measurement model can be formulated:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1\ 0\ 0 \\ 1\ 1\ 0 \\ 1\ 2\ 0 \\ 1\ 2\ 1 \\ 1\ 2\ 2 \\ 1\ 2\ 3 \end{bmatrix} * \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{bmatrix} \tag{8}$$

η_1 is the intercept, η_2 is the linear slope for the first phase, and η_3 is the linear slope for the second phase. Because of the transition point at the third panel wave, the restricted values of factor loadings of η_1 are not changing for the subsequent waves.

Growth Mixture Models for Single-Phase Data

With GMM it is possible to control for unobserved heterogeneity in the data. If the variances of the growth factors in a linear or piecewise growth curve model are not different from zero, growth mixture modeling is not necessary. The GMM extends Eqs. (1)–(3) by a categorical variable c with $k = 1, 2, \ldots, K$ classes. Assuming a two-factor growth mixture model, the following measurement and structural equations can be formulated:

$$y_{tk} = \lambda_{1tk}\eta_{1k} + \lambda_{2tk}\eta_{2k} + \epsilon_{tk} \tag{9}$$

$$\eta_{1k} = \alpha_{1k} + \zeta_{1k} \tag{10}$$

$$\eta_{2k} = \alpha_{2k} + \zeta_{2k} \tag{11}$$

Means and variances of the latent variables are estimated for each class k ($\alpha_{1k}, \alpha_{2k}, \psi_{11k}, \psi_{22k}$). The matrix Ψ_k contains the class-specific variances and covariances:

$$\Psi_k = \begin{pmatrix} \psi_{11k} \\ \psi_{21k} \ \psi_{22k} \end{pmatrix} \tag{12}$$

The so-called latent class growth analysis (LCGA; Muthén 2004) is a submodel of the GMM which gained great importance in criminological research under the name *group-based trajectory modeling* (Nagin 2005). LCGA assumes that variances and covariances of the growth factors are restricted to zero ($\Psi_k = 0$). Consequently, there are no residual terms in the structural equations of the latent variables η_1 and η_2 and therefore all class members are treated as homogenous regarding their individual developments:

$$\eta_{1k} = \alpha_{1k} \tag{13}$$

$$\eta_{2k} = \alpha_{2k} \tag{14}$$

Previous analyses of delinquent trajectories with longitudinal data show that specifications of the LCGA lead to quite reasonable substantive results (see, for instance, Kreuter & Muthén 2008). From a methodological point of view Muthén (2004, p. 350) suggests using LCGA as starting point for the analysis of trajectories, because it can be explored how many different classes might be necessary to estimate distinct developmental trends appropriately.

In most criminological studies the longitudinal response variable is a count measure (e.g., the number of convictions). Therefore, the Poisson regression model as a special case of the generalized linear model has to be used. Let y_i be the number of observed count occurrences, x_i the vector of covariates, and ν_i the expected number of counts. The number of events in an interval of a given length is Poisson distributed and the Poisson regression model can be formulated via a log link function (Hilbe 2011, p. 31):

$$Pr(y_i|x_i) = exp(-\nu_i)\nu_i^{y_i}/y_i! \tag{15}$$

with $\nu_i = exp(\alpha + x_i^{'}\beta)$. β is the vector of regression coefficients. The conditional mean function of the Poisson distribution is $E(y_i|x_i) = \nu_i$ with its equidispersion $Var(y_i|x_i) = \nu_i$. Small values of ν_i indicate the rarity of the event and the skewness of the distribution.

If the assumption of equidispersed data does not hold, the negative binomial regression model can be employed by introduction of latent heterogeneity in the conditional mean of the Poisson model (Hilbe 2011, p. 185):

$$Pr(y_i|x_i, \epsilon_i) = exp(\alpha + x_i'\beta + \epsilon_i) = h_i v_i \qquad (16)$$

where $h_i = exp(\epsilon_i)$ is assumed to have a one parameter gamma distribution, $G(\theta, \theta)$ with a mean equal to 1 and variance $\kappa = 1/\theta$. The negative binomial distribution can be obtained by integrating h_i out of the joint distribution. The conditional mean function is still $E(y_i|x_i) = v_i$ while overdispersion can be obtained from the latent heterogeneity with the variance function $Var(y_i|x_i) = v_i^2[1 + (1/\theta)]$.

Within the context of the CrimoC study previous analyses of GMM using the assumption of a negative binomial distributed variable have shown that those models have always better model fits compared to models with the assumption of a continuous or Poisson distributed variables (Reinecke & Seddig 2011, p. 432). Therefore the negative binomial distribution assumption will be used for the current analyses.

Growth Mixture Models for Multi-Phase Data

The discussed mixture models always assume that every estimated trajectory relies on longitudinal data covering a single phase of development. In case of long repeated panel designs this assumption might not be appropriate. The larger the time span of the longitudinal data, the higher is the chance that modeling of different phases is necessary to estimate the trajectories of the particular latent classes. The difference between single-phase and multi-phase data does not depend on specific features of a panel design but on whether transitions points are likely between the particular measurement occasions.

For homogenous populations piecewise growth curve models, as discussed above, are able to consider transition points. In case of unobserved heterogeneity the piecewise growth curve model can be extended to a so-called *Traditional Piecewise Growth Mixture Model* (TPGMM, Kim & Kim 2012, p. 300). TPGMM has multiple growth components and additionally one mixture component. The growth components are the same as for piecewise growth models whereas the finite mixture component is the same as for the GMM. The growth trajectories before and after a transition point are connected at the transition point. Figure 1 illustrates the model assumption: y_1-y_8 are the measures for eight panel waves, I is the intercept and $S1$ as well as $S2$ are the particular slopes. The first and second growth trajectory are connected at the transition point (e.g. t_6). c represents the mixture component, X is a time-invariant exogenous variable, U represents outcome variables. Both X and U will not be considered in the applications (cf. section "Modeling Results").

If a larger change or a discrepancy (e.g. intervention) is expected at the transition point, the TPGMM might not be sufficient to model this effect. One possible
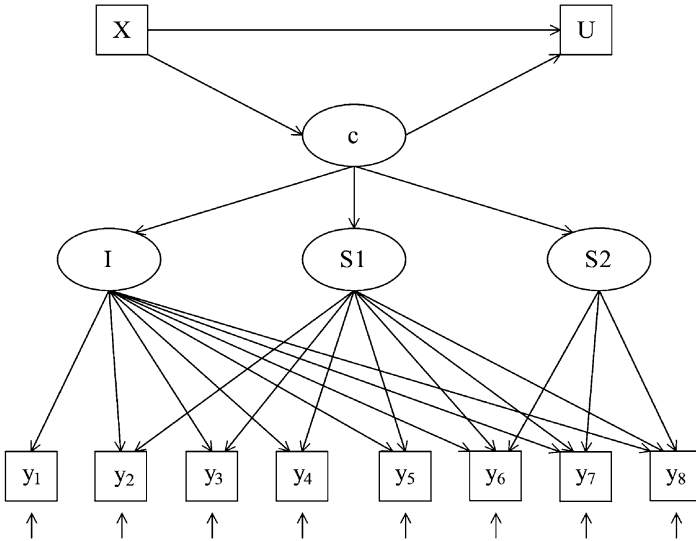
Fig. 1 Traditional Piecewise Growth Mixture Model. *Source*: Kim and Kim (2012, p. 297)

extension of the TPGMM is the so-called *Discontinuous Piecewise Growth Mixture Model* (DPGMM, Kim & Kim 2012, p. 301) in which an intercept is specified for each phase. Figure 2 shows an example with eight panel waves and a transition point between the fourth and fifth measurement: y_1–y_4 are the first-phase measures, y_5–y_8 are the second-phase measures, $I1$ is the first intercept and $S1$ is the first slope, $I2$ is the second intercept and $S2$ is the second slope. All other variables are the same as in Fig. 1. In difference to the TPGMM the trajectories of the first and the second phases are not directly connected at the transition point.

Introducing a second intercept changes the measurement part of the DPGMM compared to the TPGMM while the structural part remains the same. The measurement part of the model is given as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} 1\ 0\ 0\ 0 \\ 1\ 1\ 0\ 0 \\ 1\ 2\ 0\ 0 \\ 1\ 3\ 0\ 0 \\ 0\ 0\ 1\ 0 \\ 0\ 0\ 1\ 1 \\ 0\ 0\ 1\ 2 \\ 0\ 0\ 1\ 3 \end{bmatrix} * \begin{bmatrix} \eta_{11} \\ \eta_{21} \\ \eta_{12} \\ \eta_{22} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \end{bmatrix} \qquad (17)$$

η_{11} and η_{12} are the intercepts for the first and the second phase whereas η_{21} and η_{22} are the particular slopes. Both the TPGMM and the DPGMM assume
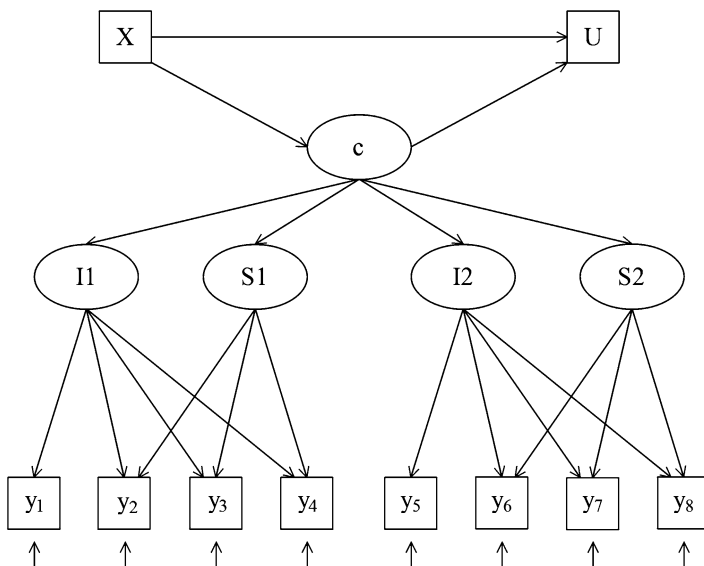
Fig. 2 Discontinuous Piecewise Growth Mixture Model

a common mixture component c for all phases. If in addition changes between the classes due to the transition between the phases have to be considered, the DPGMM can be extended to a so-called *Sequential-Process Growth Mixture Model* (SPGMM, Kim & Kim 2012, p. 303). Transition points as well as changes between latent class membership can be applied with the SPGMM. Figure 3 shows an example with the two mixture components $c1$ and $c2$. The relationship between both mixture components is specified via a transition probability matrix which contains the estimates of the probability of latent class membership of the second phase, conditional on latent class membership at the first phase. The number of intercepts and slopes and the specifications of the measurement part of the model are equal to the DPGMM.

Model Estimation and Model Evaluation

Mixture models are estimated by maximizing the log-likelihood function within the admissible range of parameter values given classes and data. The program M*plus* employs the EM-algorithm for maximization (Dempster, Laird, & Rubin 1977; Muthén & Shedden 1999). Thereby, different sets of starting values are tested for the calculation of the optimal function value and the best set is used for the estimation of the parameters. For a given solution, each individual's probability of membership in each class is estimated. Individuals can be assigned to the classes by
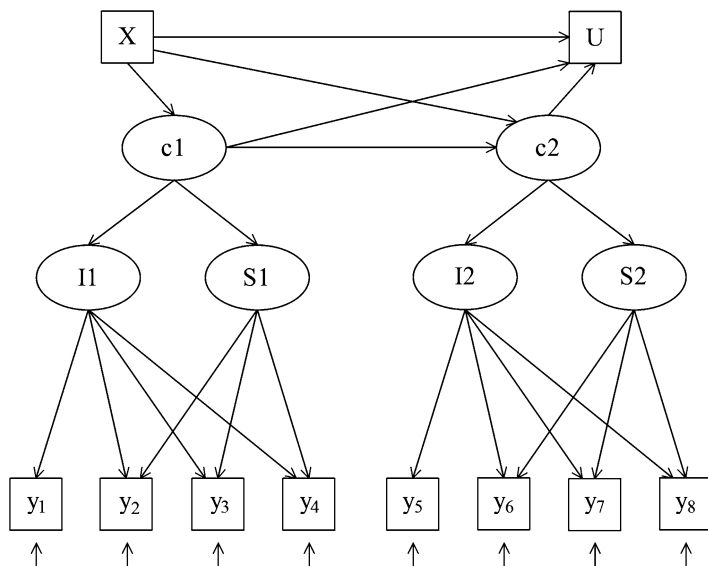
Fig. 3 Sequential-Process Growth Mixture Model

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. The number of classes has to be specified in each model variant.

Standard errors of estimates are asymptotically correct if the underlying mixture model is the true model. χ^2-differences between the particular mixture model variants, however, cannot be calculated because a k-class model is not nested within a $k + 1$-class model. Therefore, the Bayesian Information Criterion (BIC, Schwarz 1978) is used for model comparisons. Furthermore, M*plus* calculates a sample size adjusted BIC which was found to give superior performance for model selection (adj. BIC, Yang 1998). Models with the lowest BIC or adjusted BIC can be selected for further substantial interpretations. But accepting or rejecting a model on the basis of the BIC is more or less descriptive and does not imply any statistical test.

However, Lo, Mendell, and Rubin (2001) have developed a statistical test for mixture models. The so-called Lo-Mendell-Rubin likelihood test (LMR-LRT) tests a k-class model against a $k - 1$-class model. Thereby the relation of the likelihoods of a $k - 1$-class model to the ones of a k-class model is calculated. If the p-value of the test is small, the k-class model should be accepted (Reinecke 2012, p. 38). The LMR-LRT can only be calculated for GMM and TPGMM.

In addition, the entropy of a particular mixture model can be used to decide about the adequate number of classes. Entropy is a summary measure of classification quality based on the estimated posterior probabilities that ranges from zero to one:

$$E_k = 1 - \frac{\Sigma_i \Sigma_k (-\hat{p}_{ik} ln \hat{p}_{ik})}{n \, lnK} \qquad (18)$$

$\hat{p}_{ik}$ is the estimated probability for each individual i to be in class k. The closer its values are to one, the better the classification.

BIC, adjusted BIC, LMR-LRT, Entropy, and the substantive interpretability of the classes should be considered for the decision process (Muthén & Muthén 2000). In the context of multi-phase mixture models three additional aspects have to be taken into account (Kim & Kim 2012, 305f): At first, the number of latent classes in each phase should be kept as small as possible, second,for multiple latent classes nearly empty patterns can be accepted (e.g., as outliers) and finally, redundant classes should be avoided as well as classes which are misleadingly omitted. All in all, it is advisable to make the decision about the number of latent classes not only on the basis of one information source, but include various statistical and substantive arguments (see also Kim 2014).

Data, Variables, and Descriptive Statistics

The data used for the current analyses are taken from the panel survey of Duisburg which is part of the ongoing German panel study CrimoC. Duisburg is an industrial city of about 500,000 inhabitants. It is located in the western part of the Ruhr area in Germany. The sample was drawn from secondary schools in Duisburg. Eight annual panel waves have been collected between 2002 and 2009, which covers the period from early to late adolescence. The self-administered questionnaires were completed in school classes as long as the students attended the particular schools. After leaving school participants were usually contacted by mail. If repeated contacts were unsuccessful personal contacts were realized to conduct the interviews. Retention rates are between 84 and 92 % (Boers et al. 2014, p. 184).

The panel data contain individuals who participated twice in a row between 2002 and 2009 (n=3938). Table 1 gives descriptive information about each panel wave (age, sex). In the first panel wave (2002) the sample's average age is 13, in 2009 it is about 20 years. The sex ratio in each panel wave is relatively balanced although there are always more female than male participants. In 2002, for instance, 48.6 % of the respondents are male and 51.4 % female. In the subsequent panel waves, however, there are larger differences. In 2009 only 42.2 % of the respondents are male. Therefore, females are slightly overrepresented in the data.

To measure deviant and delinquent behavior, about 15 different offenses are obtained in the questionnaires of each panel wave. These offenses can be classified into property offences (burglary, theft of cars, theft out of cars; fencing, theft out of vending machines, theft of bicycles, shoplifting), violent offences (robbery, purse snatching, assault with a weapon; assault), and criminal damage offences (graffiti, scratching, other criminal damage). Concerning each of those offences, the respondents were asked whether they ever committed it (lifetime prevalence)

and whether they have committed them in the past year (annual prevalence). If they committed the particular offence in the past year, they were also asked about the frequency of their offending (annual incidence). The time-variant dependent variable of the mixture models considers the annual incidence rates which is given as the sum of the particular rates of the 15 offences.

Table 2 gives a descriptive overview of the distributions of prevalence and incidence rates. The prevalence of the self-reported criminal behavior increases in early adolescence between 2002 and 2003 and decreases later on. The peak is reached in 2003 in which the adolescents were about 14 years old. In the year 2002 nearly 31 % of the respondents reported an offence. This rate increased to 40 % in 2003 and decreased continuously down to 7 % in 2009.

Incidence mean rates are based on the number of persons who reported at least one offence in the prevalence measure. Some of the respondents gave an answer to the annual prevalences but not to the annual incidences. Therefore the numbers of persons are slightly different for each of the eight panel waves. The lower half of Table 2 shows the means of the annual incidences of the offenders. The first row of means are based on the number of valid answers in each panel wave. The second row are the means estimated via the Full Information Maximum Likelihood procedure (FIML, Enders 2010, p. 88) considering unit nonresponses in each panel

Table 1 Descriptive information about the sample

Year		2002	2003	2004	2005	2006	2007	2008	2009
n		2683	3094	3105	3140	2989	2577	2410	2299
Age		13	14	15	16	17	18	19	20
Sex	Male	48.6	48.9	47.8	48.4	45.5	43.2	43.4	42.2
	Female	51.4	51.1	52.2	51.6	54.5	56.9	56.6	57.8

Table 2 Annual prevalence and incidence rates

Year	2002	2003	2004	2005	2006	2007	2008	2009
Prevalence rates (percentages)								
n	2683	3094	3105	3140	2989	2577	2410	2299
Property	18.5	24.3	21.7	17.4	12.9	8.3	5.0	4.6
Serious	2.9	4.3	4.2	3.7	2.1	0.9	0.7	0.3
Minor	17.9	23.3	21.0	16.6	12.3	7.8	4.7	4.5
Violence	13.9	19.2	15.0	13.2	9.6	5.4	4.1	2.6
Serious	3.8	6.2	4.8	4.3	2.5	1.1	0.8	0.5
Minor	12.8	16.7	13.3	11.8	8.8	5.1	3.6	2.2
Damage	16.9	23.3	18.9	13.8	9.5	5.4	3.0	1.6
All offences	30.6	39.1	35.1	28.6	22.4	14.3	9.6	7.3
Incidence rates (means)								
n	2483	2692	2783	2838	2751	2464	2282	2205
Mean	2.78	4.83	5.28	4.26	3.32	1.74	1.70	0.77
n	3938	3938	3938	3938	3938	3938	3938	3938
Mean (FIML)	2.92	5.24	5.97	4.78	3.73	2.13	1.92	0.81

FIML Full Information Maximum Likelihood

wave. The FIML estimated means are slightly higher compared to those based on the complete cases in each panel wave which reflects a certain underreporting of the incidence rates (see also Reinecke & Weins 2013). Nevertheless, both rows of means reflect the typical development of adolescents' delinquent behavior with the peaks at age 15 (year 2004) and a continuous decline thereafter. FIML estimated means, variances, and covariances are used for the GMM in section "Modeling Results".

Modeling Results

With different slope specifications variants of the TPGMM (see Fig. 1) are firstly evaluated. One specification assumes three phases with one turning point at the second panel wave and another turning point at the sixth panel wave. The factor loadings of the intercept and the three linear slopes are restricted as follows:

$$
\begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0.5 & 0 \\
1 & 1 & 1.5 & 0 \\
1 & 1 & 2.5 & 0 \\
1 & 1 & 3.5 & 0 \\
1 & 1 & 3.5 & 20.25 \\
1 & 1 & 3.5 & 30.25
\end{pmatrix}
\tag{19}
$$

The first linear slope (second column in the matrix) specifies the first turning point. Therefore subsequent factor loadings are restricted to the value of one. The second linear slope (third column) specifies the continuous development of delinquency up to the sixth panel wave with a difference value of one. Therefore subsequent factor loadings are fixed to the value of 3.5. The third slope (fourth column) reflects a faster development by doubling the value of 3.5 with additional constants ($3.5^2 + 8 = 20.25$ and $3.5^2 + 10 = 30.25$). These fixed values were previously explored by different model specifications of the piecewise growth curve model.

Alternatively, a more parsimonious specification assumes only two slopes and a faster development of delinquency after the second wave. The factor loadings of the intercept and the two linear slopes are restricted as follows:

$$
\begin{pmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
1 & 1 & 0.25 \\
1 & 1 & 2.25 \\
1 & 1 & 6.25 \\
1 & 1 & 12.25 \\
1 & 1 & 20.25 \\
1 & 1 & 30.25
\end{pmatrix}
\tag{20}
$$

The restrictions of the factor loadings of the first slope (second column in the matrix) are equal to the previous specification in Matrix (19). The second slope (third column) specifies the continuous development of delinquency by adding the constants 2, 4, 6, 8, and 10 to the value of 0.25. So, the factor loadings of the second slope for the last two panel waves do not differ to the factor loadings from the third slope in Matrix (19). In general, restrictions of the factor loadings influence the form of the trajectory pieces, the direction of the development (increase or decrease of delinquency) can only be observed from the sign of the particular slope mean estimators (see vector α in Eq. (7)).

TPGMM are calculated from two up to eight classes. Incidence rates are treated as a negative binomial distributed count variable (cf. Eq. (16)). Intercept and slopes are specified according to LCGA, i.e., all variances and covariances of the growth curve variables are fixed to zero (cf. section "Growth Mixture Models for Single-Phase Data"). Table 3 shows the particular fit information for the TPGMM with three and two linear phases according to the specifications in Matrices (19) and (20).

All the BIC and adjusted BIC values of the models with two phases are lower than the particular models with three phases. It clearly shows that the development of delinquency can be modelled sufficiently well with two phases: one for the increase and one for the decrease. Regarding the TPGMM with two phases the p-value of the LMR-LRT shows no redundancy up to six classes.

Table 4 and Fig. 4 give an overview of the model. The largest class in this model represents a group of adolescents who were nearly not involved in delinquent behavior during the observed period (non-offenders, 49.9 %). The second largest class is characterized by a slight increase in the early adolescence and a likewise

Table 3 Model fit information of the TPGMM with three and two phases

Class	Parameter	E_k	BIC	adj. BIC	LMR-LRT	p-Value
Three phases						
2	17	0.699	46,116	46,062	2182	0.00
3	22	0.624	45,656	45,586	490	0.00
4	27	0.562	45,538	45,452	168	0.00
5	32	0.555	45,462	45,360	116	0.01
6	37	0.574	45,445	45,328	56	0.26
7	42	0.569	45,437	45,304	45	0.91
8	47	0.551	45,441	45,291	37	0.00
Two phases						
2	15	0.699	46,098	46,050	2159	0.00
3	19	0.621	45,631	45,570	485	0.00
4	23	0.612	45,470	45,397	187	0.01
5	27	0.552	45,402	45,316	99	0.01
6	31	0.546	45,374	45,276	59	0.05
7	35	0.538	45,366	45,255	39	0.42
8	39	0.543	45,368	45,244	24	0.13

Table 4 Means of the growth variables (TPGMM)

Class	Variable	Mean	Standard error	z-Value
Class 1	I	−3.180	0.413	−7.698
Non−offenders	S1	−0.833	0.542	−1.537
(n=1966, 49.9%)	S2	−0.055	0.027	−2.052
Class 2	I	1.253	0.435	2.878
Adolescence−limited	S1	0.729	0.333	2.191
(n=662, 16.8%)	S2	−0.092	0.017	−5.292
Class 3	I	0.741	0.417	1.779
Low−level−decliners	S1	0.067	0.343	0.196
(n=530, 13.5%)	S2	−1.545	0.531	−2.909
Class 4	I	−1.331	0.454	−2.931
Low−rate−offenders	S1	0.838	0.400	2.096
(n=300, 7.6%)	S2	−0.008	0.033	−0.245
Class 5	I	2.705	0.178	15.167
Persistent offenders	S1	0.827	0.216	3.825
(n=270, 6.9%)	S2	−0.033	0.013	−2.555
Class 6	I	2.231	0.315	7.076
High−level−decliners	S1	0.797	0.290	2.745
(n=210, 5.3%)	S2	−0.841	0.076	−11.090

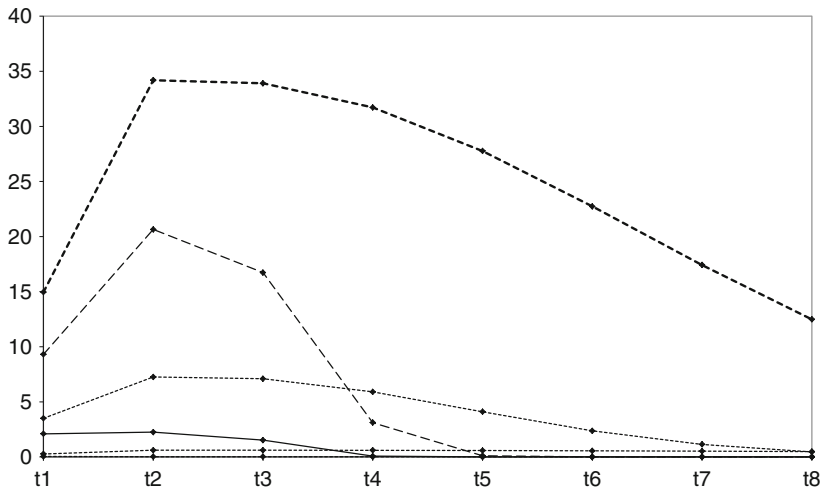*I* intercept, *S1* first slope, *S2* second slope

Fig. 4 Traditional Piecewise Growth Mixture Model with six classes. Labels of the classes (from the *bottom* to the *top*): non-offenders, low-rate-offenders, low-level-decliners, adolescence-limited-offenders, high-level-decliners, persistent offenders

slight decrease later on. Here, delinquency was limited to the period of adolescence (adolescence-limited offenders, 16.8 %). The third largest group comprises adolescents who committed crimes just in early adolescence (low-level-decliners, 13.5 %). The following class is a group of adolescents who reported only a few crimes during the observed period (low-rate-offenders, 7.6 %). Only a small proportion of the adolescents can be classified as persistent offenders with a large incidence rate (6.9 %). A likewise small proportion is characterized by a high crime rate in early adolescence and a low crime rate later on (high-level-decliners, 5.3 %).

This type of growth trajectory, however, can distort real growth patterns in data, when a more dynamic change or a discrepancy is expected at the transition point.

The TPGMM, however, "can distort real growth patterns in data, when a more dynamic change or a discrepancy is expected at the transition point" (Kim & Kim 2012, p. 300). The DPGMM (see Fig. 2) contains intercepts for each phase. For the substantive application a DPGMM would be assumed to have one intercept and slope for the increase of delinquency as well as one intercept and slope for the decrease of delinquency. With this model a larger discrepancy at the transition is expected which means a sufficient discontinuity between the phases. However, previous analyses with the CrimoC panel data did not support a discontinuity of the developmental process and therefore the specification of a DPGMM was rejected.

As described in section "Growth Mixture Models for Multi-Phase Data" the SPGMM extends the DPGMM by additional latent class variables assuming that the number of classes can change between the phases. According to the results of the DPGMM and in difference to Fig. 3 we do not assume two but only one intercept for the phases. In addition, it is proposed that the number of classes will decrease over time. Substantively this means that a larger unobserved heterogeneity of the trajectories is expected in early adolescence compared to late adolescence. With increasing age and increasing desistance from crime a smaller unobserved heterogeneity is expected. Similar to the TPGMM, the models are tested with three and two phases. According to the assumption of decreasing heterogeneity the number of classes is always higher in the first phase compared to the subsequent phases. Table 5 shows the model fit information for the calculated SPGMMs. Model selection is limited to the BIC and adjusted BIC (LMR-LRT is not calculated in M*plus* when different class patterns are specified). Similar to the TPGMM, results show that two phases are sufficient. The model with three classes in the first phase and two classes in the second phase can be selected for further interpretations.

The combination of the first and the second phase leads to a six-class pattern with different combinations of classes (see Table 6):

1. Class pattern 1 1: 12.3 % of the adolescents change from low-rate-offenders in the early adolescence to non-offenders later on.
2. Class pattern 1 2: 4.1 % of the adolescents are characterized by a high and increasing delinquency rate in early adolescence and a declining delinquency rate later on.
3. Class pattern 2 1: Nearly half of the adolescents are characterized as non-offenders in both phases.

Table 5 Model fit information of the SPGMM with three and two phases

Class			Parameter	E_k	BIC	adj. BIC
Three phases						
1. Phase	2. Phase	3. Phase				
2	1	1	17	0.614	46,116	46,062
2	2	1	26	0.563	45,532	45,450
3	2	1	35	0.552	45,428	45,317
Two phases						
2	1	–	15	0.699	46,098	46,050
2	2	–	19	0.621	45,631	45,570
3	1	–	22	0.614	45,477	45,407
3	2	–	29	0.538	45,365	45,272

Table 6 Number and proportion of persons in the class patterns (SPGMM)

Class	Class pattern	n	%
1	1 1: low-rate-offenders → non-offenders	484	12.3
2	1 2: high starters → decliners	162	4.1
3	2 1: non-offenders	1888	47.9
4	2 2: early increasers → decreasers	759	19.3
5	3 1: low-rate-offenders	415	10.5
6	3 2: high starters → persisters	230	5.8
Latent variable	Composition	n	%
C1	1: early starters/high starters (484 + 162)	646	16.4
(Phase 1)	2: non-offenders/early increasers (1888 + 759)	2647	67.2
	3: low rate offenders/high starters (415 + 230)	645	16.4
C2	1: non-offenders/low-rate-offenders (1888 + 415)	2787	70.8
(Phase 2)	2: decliners/decreasers/persisters (162 + 759 + 230)	1151	29.2

4. Class pattern 2 2: 19.3 % of the adolescents show a slight increase in early adolescence and a slight decrease later on.
5. Class pattern 3 1: 10.5 % of the adolescents are characterized as low-rate-offenders in both phases.
6. Class pattern 3 2: 5.8 % of the adolescents show persistent delinquency on a high level with a decreasing tendency in late adolescence.

The first phase is characterized by three classes. The first one comprises adolescents who started to behave delinquently early in the adolescence and partly on a high level (16.4 %). The second class encompasses non-offenders and adolescents whose crime rate increases slightly on a low level (67.2 %). Finally, low rate offenders and high starters can be found in the third class. The second phase comprises two classes. The first of them encompasses non- and low-rate offenders, the second one adolescents with decreasing delinquency (see Fig. 5).
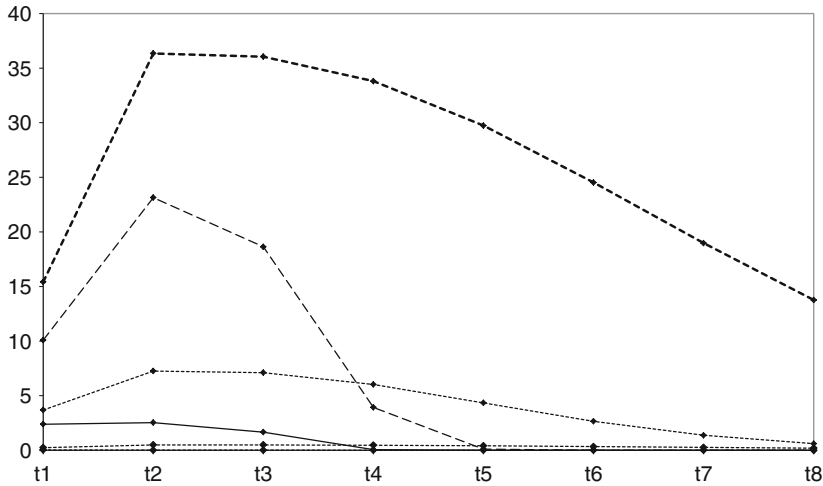
Fig. 5 Sequential-Process Growth Mixture Model. Labels of the classes (from the *bottom* to the *top*): non-offenders, low-rate-offenders, low-rate-offenders → non-offenders, early increasers → decreasers, high-starters → decliners, high-starters → persisters

With the two latent class variables C1 and C2 the SPGMM is able to estimate transition probabilities. With a probability of 66 % it is more likely for adolescents to stay as or to become non- or low-rate-offenders during the life course than to still act delinquent in late adolescence. Quite a few adolescents, however, commit crimes in late adolescence as well.

One possibility to compare and validate the results of different mixture model specifications is to look at the bivariate table with the particular proportions for the latent classes based on their most likely latent class membership. The estimated TPGMM contains six classes and one latent class variable, the estimated SPGMM also contains six classes which can be differentiated into different class patterns. Table 7 gives the result of the crosstabulation between the class distribution of the TPGMM and the SPGMM. Most of the individuals in pattern 1 1 of the SPGMM belong to the third class of the TPGMM (low-level decliners, 96.07 %), nearly all individuals in pattern 1 2 of the SPGMM belong to the sixth class of the TPGMM (high-level decliners, 99.38 %). Differences between these two patterns (SPGMM) or classes (TPGMM) refer only to the level of delinquency, both patterns or classes are characterized by processes of desistance.

Non-offenders in pattern 2 1 of the SPGMM are 100 % part of class 1 of the TPGMM. Pattern 2 2 of the SPGMM is characterized by processes of early increasing and later declining delinquency. Eighty-six percent of these individuals belong to class 2 of the TPGMM (adolescent-limited offenders). The rest of pattern 2 2 is distributed across the other classes of the TPGMM. The lowest congruence to class 4 of the TPGMM has pattern 3 1 of the SPGMM (low-rate offenders). Only 68 % of the individuals are in the particular cell of the cross-table. Nearly 17 %

Table 7 Individual latent class membership of the TPGMM (row) and SPGMM (column)

Class	1 1 low-rate-offenders → non-offenders	1 2 high starters → decliners	2 1 non-offenders	2 2 early increasers → decreasers	3 1 low-rate-offenders	3 2 high starters → persisters	Σ
1 Non-offenders	9	0	1888	0	69	0	1966
	0.46	0.00	96.03	0.00	3.51	0.00	100.00
	1.86	0.00	100.00	0.00	16.63	0.00	49.92
2 Adolescence-limited	2	0	0	653	7	0	662
	0.30	0.00	0.00	98.64	1.06	0.00	100.00
	0.41	0.00	0.00	86.03	1.69	0.00	16.81
3 Low-level-decliners	465	0	0	23	42	0	530
	87.74	0.00	0.00	4.34	7.92	0.00	100.00
	96.07	0.00	0.00	3.03	10.12	0.00	13.46
4 Low-rate-offenders	0	0	0	16	284	0	300
	0.00	0.00	0.00	5.33	94.67	0.00	100.00
	0.00	0.00	0.00	2.11	68.43	0.00	7.62
5 Persistent offenders	0	1	0	39	0	230	270
	0.00	0.37	0.00	14.44	0.00	85.19	100.00
	0.00	0.62	0.00	5.14	0.00	100.00	6.86
6 High-level-decliners	8	161	0	28	13	0	210
	3.81	76.67	0.00	13.33	6.19	0.00	100.00
	1.65	99.38	0.00	3.69	3.13	0.00	5.33
Σ	484	162	1888	759	415	230	3.938
	12.29	4.11	47.94	19.27	10.54	5.84	100.00
	100.00	100.00	100.00	100.00	100.00	100.00	100.00

of the pattern belongs to the non-offender class 1 and about 10 % to the low-level declining class 3 of the TPGMM. Similar to the class of non-offenders pattern 3 2 (persistent offenders) of the SPGMM are 100 % part of class 5 of the TPGMM. In total, the crosstabulation of both class memberships confirms the stability of the latent class distributions although the specifications of TPGMM and SPGMM are different. Non-offenders and low-rate offenders have overlaps in their particular developments and therefore their assignments can differ between the models. This has been observed in previous applications of GMM with criminological panel data (see, for example, Mariotti & Reinecke 2010; Piquero 2008; Reinecke & Seddig 2011).

Conclusion

This study has shown that with an increasing number of panel waves unobserved heterogeneity of developmental processes results not only from a mix of these developments but also from multiple phases. In difference to the TPGMM, the SPGMM has separate mixture parts with a latent class variable in each phase. Whereas the TPGMM has only one intercept over multiple phases, the DPGMM and SPGMM specify separate intercepts as well as separate slopes. Kim and Kim (2012) showed how growth and mixture models can be extended to more complex and flexible stage-sequential growth mixture models within the structural equation modeling framework. Their model applications contain continuous data related to smoking behavior. In the present study the observed variable was treated as a count variable with overdispersion. Therefore piecewise and stage-sequential growth mixture models have been applied with the specification of a negative binomial distributed variable. But all the analyses are limited to the LCGA specification meaning that no variances and covariances were estimated for the growth curve variables within classes.

   With eight panel waves of self-reported delinquency obtained from the CrimoC study (Boers et al., 2014) separate intercepts could not be detected and identified while separate growth components reflect increase and decrease of delinquency through the period of adolescence and young adulthood. If only one intercept is required, the specification of the DPGMM collapses to the TPGMM. One possible explanation is that the CrimoC study contains no experimental intervention and therefore the different trajectory pieces do not reflect phases of discontinuity.

   Starting with the single-phase TPGMM six distinct classes of delinquent developments could be identified: non-offenders who were nearly not involved in delinquent behavior at all, adolescent-limited offenders with the typical development of the age-crime curve, low-level-decliners who limited their delinquency in early adolescence, low-rate-offenders who reported only a few crimes during the panel study, persistent offenders with the largest incidence rate compared to the other classes and high-level decliners with a high crime rate in early adolescence and a declining tendency later on. A specification with six classes could also be verified

with the multiple-phase SPGMM. Different models with two and three phases were tested and compared. The first phase can be characterized by the development of delinquency in early adolescence, the second phase by the development in late adolescence. A possible third phase belongs to the period of young adulthood which might be detected with further panel waves. The specification of the different SPGMM variants assume always a higher number of classes in the first phase compared to the second or third phase. Heterogeneity of the development of delinquency is expected to be higher in the first panel waves and decreases thereafter. On the average this assumption was confirmed. The number of offenses decreases over time and the development of delinquency tends to be homogenized. Two class patterns of the final SPGMM are expected to be stable across the phases: the non-offenders and the low-rate offenders. One pattern shows the transition from low-rate to non-offending, two patterns show the transition from high starters to decliners or persisters and another pattern is characterized by the transition of early increasing to later decreasing delinquency. Transition parameters between the phases show that the probability to stay as or to become a non- or low-rate-offender is much higher than to persist as a delinquent persons during the life course. The crosstable of the most likely latent class memberships of the TPGMM and the SPGMM reflects the stability of the classification and serve as a proof of quality for the substantive interpretations.

Although the applications of the single and multi-phase mixture models is very useful for the longitudinal criminological research technique in various fields, some unresolved issues have to be mentioned. The complexity of the models requires not only large sample sizes but also a large number of starting values. In the initial stage, 500 random sets of starting values were generated and optimized for each of the sets. The ending values of 20 optimizations with the highest log-likelihoods were used as starting values in the final stage. With the assumption of a negative binomial distribution stable results could only be obtained with the LCGA specification. Evaluation of model fit is not the same for single-phase and multiple-phase mixture models. The LMR-LRT is only available for models with one latent class variable while the statistical evaluation of multiple-phase models is limited to descriptive information criteria with preference to the adjusted BIC (Kim 2014). In addition, the large number of zeros in the incidence rates can be accounted by an inflation part of the particular mixture model (Reinecke & Seddig 2011). This extension has to be studied in future applications of stage-sequential growth mixture models.

References

Boers, K., Reinecke, J., Daniel, A., Kanz, K., Schulte, P., Seddig, D., et al. (2014). Vom Jugend-zum frühen Erwach-senenalter. Delinquenzverläufe und Erklärungszusammenhänge in der Ver-laufsstudie Kriminalität in der modernen Stadt. *Monatsschrift für Kriminologie und Strafrechtsreform, 3*, 183–202.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach. Wiley series on probability and mathematical statistics.* New Jersey: Wiley.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplète data via the EM algorithm. *Journal of the Royal Statistical Society, 39*, 1–38.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Mahwah: Lawrence Erlbaum.

D'Unger, A., Land, K. C., McCall, P. L., & 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.

Enders, C. K. (2010). *Applied missing data analysis.* New York: Guilford Press.

Farrington, D. P., & West, D. J. (1990). The Cambridge study in delinquent development: A longterm follow-up of 411 London maies. In: H. J. Kerner & G. Kaiser (Eds.), *Kriminalität: Persönlichkeit, Lebensgeschichte und Verhalten* (pp. 115–138). Berlin: Springer.

Hilbe, J. M. (2011). *Negative binominal regression* (Vol. 2). Cambridge: Cambridge University Press.

Kim, S.-Y. (2014). Determining the number of latent classes in single- and multiphase growth mixture models. *Structural Equation Modeling: A Multidisciplinary Journal, 21*, 263–279.

Kim, S.-Y., & Kim, J.-S. (2012). Investigating stage-sequential growth mixture models with multiphase longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 19*, 293–319.

Kreuter, F., & Muthén, B. O. (2008). Analyzing criminal trajectory profiles: Bridging multilevel and groupbased approaches using growth mixture modelling. *Journal of Quantitative Criminology, 24*, 1–31.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767–778.

Loeber, R., & LeBlanc, M. (1990). Toward a developmental criminology. In T. Michael & M. Norval. (Eds.), *Crime and justice* (Vol. 12, pp. 275–473). Chicago: University of Chicago Press.

Mariotti, L., & Reinecke, J. (2010). *Wachstums- und Mischverteilungsmodelle unter Berücksichtigung unbeobachteter Heterogenität: Empirische Analysen zum delinquenten Verhalten Jugendlicher in Duisburg.* Technical Report Sozialwis- senschaftliche Forschungsdokumentation 21. Munster: Institut für sozialwis- senschaftliche Forschung e.V.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58*(1), 110–133.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674–701.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behavior.* Cambridge: Cambridge University Press.

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* (pp. 345–368). Thousand Oaks: Sage.

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable- centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism, Clinical and Experimental Research, 24*, 882–891.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55*, 463–469.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods, 4*, 139–157.

Nagin, D. S. (2005). *Group-based modeling of development.* Cambridge, MA: Harvard University Press.

Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. *Criminology, 31*, 327–362.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., et al. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology, 20*, 673–716.

Piquero, A. R. (2008). Taking stock of developmental trajectories of criminal activity over the life course. In A. M. Liberman. (Ed.), *The long view of crime: A synthesis of longitudinal research* (pp. 23–78). New York: Springer.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks: Sage.

Reinecke, J. (2012). *Wachstumsmodelle.* Sozialwissenschaftliche Forschungsmeth- oden Band 3. München und Mering: Rainer Hampp Verlag.

Reinecke, J., & Seddig, D. (2011). Growth mixture models in longitudinal research. *Advances in Statistical Analysis, 95*, 415–434.

Reinecke, J., & Weins, C. (2013). The development of delinquency during adolescence: A comparison of missing data techniques. *Quality & Quantity, 47*(6), 3319–3334.

Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turningpoints through life.* Cambridge: Harvard University Press.

Sampson, R. J., & Laub, J. H. (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. In T. P. Thornberry. (Ed.), *Developmental theories of crime and delinquency. Advances in criminological theory* (Vol. 7, pp. 133–161). New Brundwick, London: Transaction Publishers.

Sampson, R. J., & Laub, J. H. (2003). Life-course desisters? Trajectories of crime among delinquent boys followed age 70. *Criminology, 41*, 555–592.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Thornberry, T. P. (2005). Explaining multiple patterns of offending across the life course and across generations". In R. J. Sampson & J. H. Laub. (Eds.), *Developmental criminology and its discontents. The annals of the American Academy of Political and Social Science* (Vol. 602, pp. 156–195). Thousand Oaks: Sage.

Tracy, P. E., Wolfgang, M. E., & Figlio, R. M. (1990). *Delinquency in two birth cohorts.* New York: Plenum.

Tremblay, R. E., Desmarais-Gervais, L., Gagnon, C., & 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.

Yang, C. C. (1998). Finite mixture model selection with psychometric applications. Unpublished doctoral dissertation. University of Groningen.