



Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond

Elena A. Erosheva,^{1,2} Ross L. Matsueda,³
and Donatello Telesca⁴

¹Department of Statistics, ²School of Social Work, and ³Department of Sociology, University of Washington, Seattle, Washington 98195; email: elena@stat.washington.edu, matsueda@uw.edu

⁴Department of Biostatistics, University of California, Los Angeles, California 90095; email: dtelesca@ucla.edu

Annu. Rev. Stat. Appl. 2014. 1:11.1–11.32

The *Annual Review of Statistics and Its Application* is online at statistics.annualreviews.org

This article's doi:
10.1146/annurev-statistics-022513-115701

Copyright © 2014 by Annual Reviews.
All rights reserved

Keywords

group-based modeling, growth curve mixture modeling, latent variables, longitudinal data, unimodal curve registration

Abstract

Studies of human development require longitudinal data analysis methods that describe within- and between-individual variation in developmental and behavioral trajectories. This article reviews life-course data analysis methods for modeling these trajectories, as well as their application in studies of antisocial behavior and of crime in childhood, in adolescence, and throughout life. We set the stage by introducing growth curve (hierarchical linear) models. We focus our review on finite mixture models for life-course data, known as group-based trajectory and growth mixture models. We then discuss how these models are applied within criminology and developmental psychology, recent controversies over their substantive use and interpretation, and important issues of statistical practice and the challenges they raise. Building on the critical literature, we offer several recommendations for the applied users of the models. Finally, we present the most recent method of examining behavioral trajectories in criminology, the unimodal curve registration (UCR) approach. We briefly contrast the UCR model with growth curve and finite mixture models for life-course data analysis.

1. INTRODUCTION

The task of describing trajectories of human behavior over time is central to many problems in the social sciences. This task is particularly significant in the sciences that view human behavior from a life-course or developmental perspective and requires longitudinal data (repeated observations of a collection of subjects over a particular temporal interval). Whether the interval is short or long relative to the life span, one major goal of the analysis is always the same: to explain the within- and between-individual variability in the outcome. Although this general goal sounds similar to the classical aims of longitudinal data analysis in biostatistics (for a recent review, see Gibbons et al. 2010), the high heterogeneity in observed life-course data, together with the availability of developmental theories, has resulted in the creation of specialized methods for analysis of life-course data that go beyond traditional longitudinal analysis.

In criminology, an adequate description of offending trajectories across age is necessary for understanding differences in criminal careers (Blumstein & Cohen 1987); for estimating features of age-crime curves (Hirschi & Gottfredson 1983), such as age at onset, age at peak offending, and age at desistance; and ultimately, for explaining differences in age-crime curves using developmental or life-course theories. Early research on criminal careers and age-crime trajectories was purely descriptive; it followed the pioneering work of Wolfgang et al. (1972), who examined age at onset, criminal career length, high-rate offenders, and patterns of desistance. Attention then switched to modeling average age-crime curves, and the techniques for estimation of the population parameters evolved from a reliance on cross-sectional statistics (Hirschi & Gottfredson 1983) to the use of parametric functions (Britt 1992, Greenberg 1991). Following the introduction of random effects and mixed effects modeling for longitudinal data (Laird & Ware 1982), life-course criminology researchers have turned to model-based statistical approaches (Raudenbush 2001). Most recently, developmental criminologists have come to rely on group-based trajectory (Nagin 2005) and growth mixture modeling (GMM) (Muthén & Shedden 1999; see Piquero 2008).

The main feature that distinguishes group-based trajectory and GMM methods from other hierarchical modeling approaches is their ability to accommodate population subgroups with qualitatively different patterns of development over time, such as so-called different types of offenders in criminology. These methods and their variations have appeared in the literature under many names. Thus, group-based trajectory models may be described as a semiparametric group-based approach (Nagin & Tremblay 1999), a nonparametric, mixed model (Nagin & Land 1993), latent class growth analysis (Muthén 2004), or latent class trajectory analysis.

Despite controversy about the substantive use and interpretation of group-based trajectory and GMMs, particularly in criminology and psychology (e.g., Bauer & Curran 2003a, Nagin & Tremblay 2005b, Sampson & Laub 2003, Sampson & Laub 2005, Sterba et al. 2012), the number of publications that employ these models has been growing rapidly since the early 2000s. **Figure 1** shows yearly counts for the numbers of articles that contain the exact phrases “group-based trajectory” and “growth mixture model” as estimated by Google Scholar. Most of these publications discuss finite mixture models applied to life-course data.

This review presents an overview of recent model-based approaches for analyzing developmental and behavioral trajectories, along with applications of these approaches. To illustrate applications of group-based trajectory models to longitudinal data, we introduce a data set that has been used in early research on life-course trajectories. We then introduce two main types of discrete mixture models for longitudinal data (group-based trajectory models and GMMs), summarize recent critical opinion papers about their application, and review empirical applications in light of these criticisms. We observe that applications in criminology predominantly identify subgroup trajectories that differ in terms of two major characteristics: the overall level of

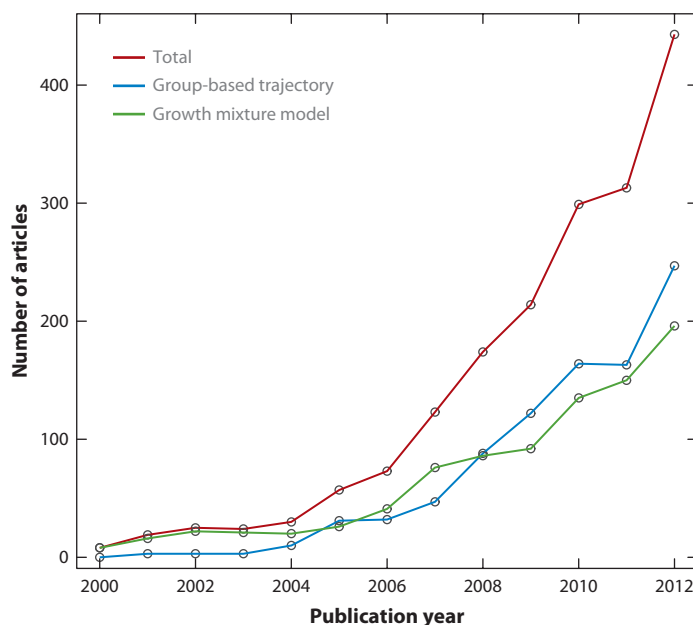


Figure 1

Google Scholar estimates for the number of articles that contain the exact phrases “group-based trajectory” (*blue*), “growth mixture model” (*green*), and their sum (*red*) by publication year.

the trajectory and the timing and duration of its peak. We then introduce an alternative model, unimodal curve registration (UCR). This model avoids a categorical mixture representation and allows one to attribute between-individual variability explicitly to differences in two sources: the level and the timing of the observed processes. Finally, we offer several recommendations for applied researchers interested in analyzing data on developmental and behavioral trajectories.

2. WORKING EXAMPLE: MONTREAL OPPOSITION DATA

For illustration purposes, consider the opposition data from the Montreal Longitudinal and Experimental Study, which began in 1984 and originally aimed to explore the development of antisocial behavior in children from kindergarten through high school (Tremblay et al. 2003). The data include opposition scores for 916 boys who were first assessed at age 6 and then assessed annually at ages 10 through 15. The opposition scores, which range from 0 to 10, were based on teacher reports of five items: does not share, demonstrates irritability, demonstrates disobedience, blames others, and shows inconsiderate behavior. These data (about 1,037 boys) were used by Jones et al. (2001) as a working example in a tutorial-style paper on group-based trajectory modeling.

The Montreal opposition data presented in **Figure 2** have several features that are characteristic of life-course or developmental data sets (compare with those reviewed in Piquero 2008). First, the opposition outcome shows a large between- and within-individual variability, which is typical for many behavioral and/or developmental outcomes. Second, many data sets include a modest number of measurement occasions, seven in this case. Third, the nine-year age span for these data is not unusual, nor is the focus on the life-course period from childhood to early adolescence.

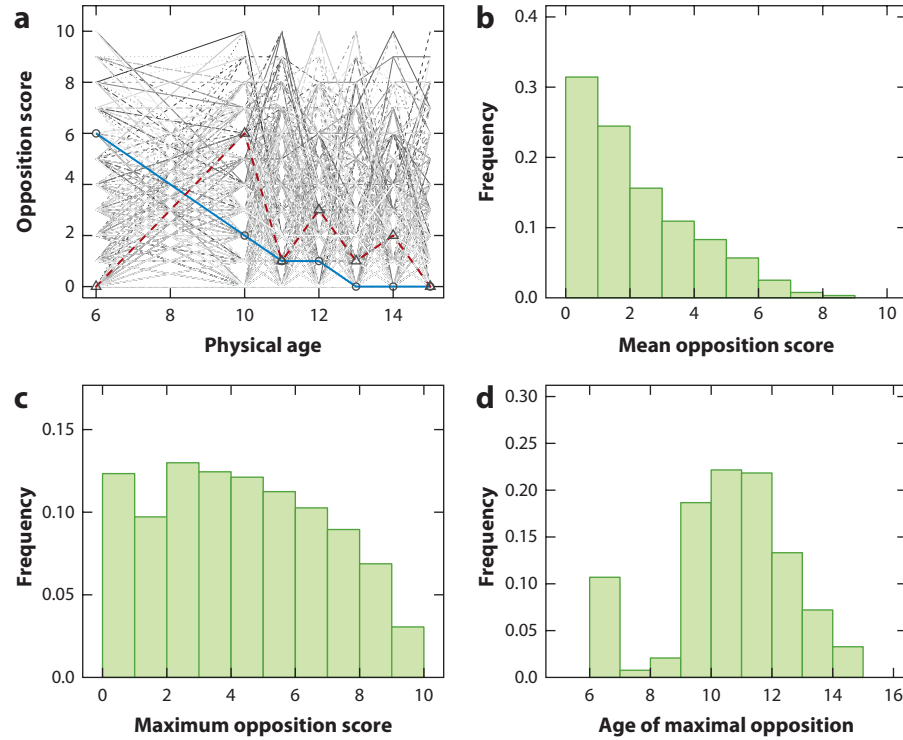


Figure 2

Opposition data. (a) Longitudinal opposition score trajectories. Two individual patterns are highlighted in color. (b) Histogram of subject-specific mean opposition scores. (c) Histogram of subject-specific maximum opposition scores. (d) Histogram of the subject-specific age at which maximal opposition is recorded.

3. TRADITIONAL MODEL-BASED APPROACHES FOR LIFE-COURSE DATA ANALYSIS

In this section, we briefly review traditional statistical modeling options for life-course data analysis. We begin by discussing growth curve models for longitudinal data, which are also known as random effects models (Laird & Ware 1982), multilevel models (Goldstein 2011), hierarchical linear models (Raudenbush & Bryk 2002), or mixed effects models (Pinheiro & Bates 2000). Growth curve models provide a basis for formulating mixture models for life-course data.

For sections 3.1 and 3.2, let $\mathbf{y}_i = (y_{i1}, \dots, y_{ij}, \dots, y_{im})'$ denote the outcome for subject i (where $i = 1, \dots, n$) over a discrete sampling design with time $\mathbf{t} = (t_1, \dots, t_j, \dots, t_m)'$. For the Montreal data, y_{ij} is the observed opposition score for subject i at time t_j . (Note that this notation also applies elsewhere in this article and is reintroduced as needed.)

3.1. Growth Curve Models

Consider a simple variation of a hierarchical linear growth model with individual-specific terms for the intercept and slope:

$$y_{ij} = \beta_0 + v_{0i} + \beta_1 t_{ij} + v_{1i} t_{ij} + e_{ij}. \quad 1.$$

Here, $\mathbf{v}_i = (v_{i0}, v_{i1})$ is a vector of random effects, the individual-specific random intercept and slope coefficients, which are assumed to be normally distributed with mean zero and variance-covariance matrix Σ_v . Random effects represent individual deviations from the mean intercept and slope in the population $\beta = (\beta_0, \beta_1)$. The residual variance $e_{ij} \sim N(0, \sigma_e^2)$ is typically assumed to be constant over time and across individuals, but both assumptions can be relaxed. The linear growth hierarchical model can be extended to incorporate polynomial or semiparametric specifications for time dependence in the mean function and in the random effects.

Let $f_y(\mathbf{t}; \beta, \Sigma_v, \sigma_e)$ denote the appropriate functional specification for the hierarchical linear growth of the outcome y across time \mathbf{t} that includes assumptions about variability of the random effects and the error term. The hierarchical linear growth model with a functional specification given above can be generalized to account for nonnormal outcomes, e.g., Poisson or negative binomial outcomes. Because generalized linear growth models impose structural variability on the error, the residual error term is no longer included as an explicit addition to $f_y(\mathbf{t}; \beta, \Sigma_v)$; rather, dispersion and possibly overdispersion parameters are defined following specific likelihood or quasi-likelihood formulations. In addition, for any type of outcome, covariates can be included to further explain variability at each level of the hierarchy. The reader is referred to Gibbons et al. (2010) for a review of recent advances in longitudinal data analysis.

3.2. Mixture Models for Life-Course Data Analysis

Instead of using a common functional specification for the hierarchical linear growth $f_y(\mathbf{t}; \beta, \Sigma_v, \sigma_e)$, researchers may assume that the data come from a finite mixture for which each component has its own specification for the growth function. The mixture components in this case correspond to unobserved groups or subpopulations with different patterns of developmental trajectories. Thus, the marginal density of the data can be written as a sum,

$$\sum_{k=1}^K p_k f_y(\mathbf{t}; \beta_k, \Sigma_v(k), \sigma_{e(k)}), \quad 2.$$

where K is the number of groups or subpopulations; p_k is the proportion of the population in group k ; and $\Sigma_v(k)$ and $\sigma_{e(k)}$ are variance specifications for k th group random effects and the random error term, respectively. This notation allows the variances of the random effects and the error term to differ by group in the most general case.

Both growth mixture (Muthén 2004, Muthén & Asparouhov 2008, Muthén & Shedden 1999) and group-based trajectory models (Nagin 2005, Nagin & Tremblay 1999, Roeder et al. 1999) follow the finite mixture specification in Equation 2, but these models differ in their assumptions regarding random effects. GMM allows for random effects specification within groups and estimates the random effects variance-covariance matrices $\Sigma_v(k)$, which can be constrained to be the same across groups. Group-based trajectory modeling is a special case of GMM that estimates only residual variability, denoted σ_e , and assumes no additional within-group random effects. Essentially, group-based trajectory modeling assumes that all entries in the matrices $\Sigma_v(k)$ are zero. Both models typically assume that the residual error variance is common across time and across groups, although these assumptions can be relaxed. Both methods also allow for (a) outcome distributions other than the normal distribution to be specified (Muthén 2004, Nagin 2005) and (b) covariate effects and the use of multinomial logistic regression to predict class assignment on the basis of covariates. Each type of mixture modeling for life-course data has been implemented in mainstream software and relies on maximum likelihood for parameter estimation.

Proc Traj, a free add-on procedure for SAS, implements group-based trajectory models using a generalized quasi-Newton procedure (Jones et al. 2001, Jones & Nagin 2007). Outcome distribution choices in this package include censored (or regular) normal, zero-inflated (or regular) Poisson, and Bernoulli (logit link). Mean group trajectories are allowed to be linear, quadratic, or cubic polynomials. Jones & Nagin (2007) describe specifications for various extensions of group-based modeling with Proc Traj. They provide several examples, including estimation of multitrajectory models that link trajectories among several outcomes and adjustments for survey weights and exposure time in the zero-inflated Poisson.

GMMs have also been implemented in the Mplus program (Muthén & Muthén 2012). This framework allows group-based trajectory models to be estimated as a special case that assumes no structural variability in random effects within population groups, referred to as latent class growth analysis. Maximum likelihood estimation is done using the accelerated EM algorithm, and three different integration methods may be used (Muthén & Asparouhov 2008). Mean group trajectories can be specified as polynomial, exponential, logarithmic, and piecewise linear (Muthén & Muthén 2005). Mplus allows model specification for continuous, binary, ordered polytomous, and zero-inflated Poisson models, and it offers options for treatment of missing and clustered data and for very flexible specification of covariance structures, including measurement errors. Thus, many model variations that can be estimated in Mplus, including the analysis of multiple indicators and two-part or semicontinuous modeling. The GMM framework also includes distal outcome analyses that account for uncertainty associated with group assignments. Generalized GMMs allow for nonparametric specification of the distribution of random effects (Muthén & Asparouhov 2008).

4. RECENT ADVANCES, REVIEWS, TUTORIALS, AND OPINION PIECES

4.1. Published Reviews, Tutorials, and Methodology Papers

In the past few years, several manuscripts have addressed various methodological and applied considerations of mixture modeling for life-course data analysis. Muthén (2004) gave an overview of GMM and related techniques for longitudinal data from the perspective of the general latent variable model. Within this framework, mixture models for life-course data can be thought of as latent variable models that combine continuous and categorical latent variables. Nagin (2005) presented a book-length treatment of group-based trajectory modeling that includes practical recommendations for estimation and model selection. Muthén & Asparouhov (2008) reviewed GMMs as a means to accommodate non-Gaussian individual-specific deviations from average trajectories; discussed maximum likelihood estimation; and provided examples from criminology, mental health, and education. Nagin & Odgers (2010a) reviewed applications of group-based trajectory modeling in clinical research. Pickles & Croudace (2010) reviewed mixture models for longitudinal outcomes with a focus on applications in clinical trials that could allow researchers to identify subgroups of patients that differ in treatment responsiveness. Nagin & Odgers (2010b) pointed to a rapid increase in applications of group-based trajectory analysis across diverse settings, including studies of the etiology and developmental course(s) of different types of disorders, studies of treatment responses in clinical trials, epidemiological observational studies, and studies of disability trajectories in elderly populations.

Finite mixture models have long been used to model complex distributions by appropriately choosing components to represent local areas of support of the true distribution (McLachlan & Peel 2000). Finite mixture models that use latent groups as points or, more appropriately, trajectories of support for life-course data are no exception. Several studies based on simulated data

found that group-based trajectory models and GMMs with relatively few groups can approximate complex distributions of trajectories reasonably well (Brame et al. 2006; Muthén & Asparouhov 2008, pp. 158–61; Nagin 2005, ch. 3; Sterba et al. 2012). Sterba et al. (2012) used simulations to explore the conditions under which the error of approximating a continuous distribution using group-based trajectory models is smaller than the error from a hierarchical generalized linear model (HGLM) with a nonnormal distribution of random effects. They found that the HGLMs do well unless the random effect distributions depart radically from normality. They also showed that mixture approximations perform well unless (a) the sample is small, (b) the random effects are numerous, and (c) random effects are uncorrelated (Sterba et al. 2012, p. 621).

Several applied works have considered the question of choice between group-based and GMM methods, the key issue in such works being whether or not random effect parameters should be included to capture within-group variability around group means (e.g., Saunders 2010). Although some researchers argue for employing theoretical considerations about structural variability around group means [for example, Nagin (2005) proposes that the introduction of within-group variation can cloud the meaning of latent groups], most approach the question from the perspective of model choice. Thus, Kreuter & Muthén (2008) carried out an empirical comparison of four models, a growth curve model and three mixture models, with varying assumptions regarding random effects. Their study used two criminological data sets, the well-known Cambridge study and the Philadelphia birth cohort study, and provided step-by-step instructions for the respective analyses. They found that for the Cambridge data, neither the shape nor the proportions of latent groups varied substantially across the mixture modeling alternatives. For the Philadelphia data, the GMMs seemed to subsume the variation across additional groups estimated with group-based modeling methods, resulting in an improved model fit. However, for a smaller subsample of $n = 500$, a conventional growth curve model with random intercept and slope provided the best fit as judged by the Bayesian information criterion (BIC) (Kass & Raftery 1995, Schwarz 1978).

Although the applied literature on growth mixture and group-based trajectory models is rich, publications in the statistical literature are relatively scarce. Statistical methodology papers include the seminal works of Muthén & Shedden (1999) and Roeder et al. (1999). Recently, Wang et al. (2005) focused on graphical diagnostics for GMMs: They examined residuals to detect misspecifications regarding covariance structure, growth trajectory means, and the number of growth classes. Tarpey et al. (2010) developed a related clustering method for linear mixed effects models and compared their methodology to growth-curve mixture modeling.

Haviland & Nagin (2005) approached causal inference with growth mixture and group-based models by developing a group-based propensity score methodology for estimating treatment impact on life-course trajectories. Using GMM, Muthén et al. (2002) assessed intervention effects in randomized trials, Muthén & Brown (2009) focused on estimating drug effects in the presence of a placebo response, and Sobel & Muthén (2012) focused on estimating the average treatment effect among compliant patients.

In addition, several authors developed Bayesian approaches for mixture models of life-course data (Elliott et al. 2005, Leiby et al. 2009, Lu et al. 2011, Neelon et al. 2011, Rolfe et al. 2011). Rolfe et al. (2011) suggest the use of latent basis trajectory specifications in a Bayesian framework. Elliott et al. (2005), Leiby et al. (2009), and Neelon et al. (2011) developed a methodology for multivariate longitudinal outcomes.

4.2. Critical Opinion Pieces

In this section, we summarize the critical evaluations of empirical applications of latent trajectory mixture models to life-course data. This literature raises important substantive issues about the

use of these models in criminology and psychology. We discuss whether these critical comments have influenced empirical applications in Section 8.2.

The field of criminology produced a series of opinion papers that questioned the scientific merits of latent mixture modeling for life-course data in general and for group-based trajectory modeling in particular (e.g., Brame et al. 2012; Nagin & Tremblay 2005a,b; Sampson & Laub 2003, 2005; Skardhamar 2010). Eggleston et al. (2004) conducted a sensitivity analysis showing that the number and mean shapes of latent classes of trajectories are not robust to either the length of the follow-up period (longer age ranges result in more groups) or failure to account for incarceration and mortality. Using the classic Glueck data on delinquent boys extended to age 70, Sampson & Laub (2003) found that when the population is subdivided into groups prospectively on the basis of levels of childhood risk factors, the predicted age-crime curves retained the same shape: a precipitous increase to the peak age of offending, followed by a slow decline until age 70. The curves differed only in level. Sampson & Laub (2003) then used a group-based trajectory model to identify latent groups of trajectories for which each group mean trajectory followed a strong unimodal shape, in contrast to the life-course-persistent group predicted by Moffitt (1993). Sampson & Laub (2003) concluded that the group-based trajectory models have seduced some criminologists by giving the appearance that there exist distinct and predictable groupings that may be amenable to direct policy intervention, such as early childhood intervention and selective incapacitation. In response, Nagin & Tremblay (2005b) argued that although group-based trajectory models are wholly appropriate when there is strong theoretical reason to believe that individual trajectories belong to trajectory groups, such a situation is unlikely to hold when examining behavioral trajectories across the life span.

Most recently, Skardhamar (2010) used a simulation study to argue that fitting group-based mixture models cannot provide evidence either for or against the existence of distinct groups because similar distributions of trajectories may result from data-generating mechanisms other than finite mixtures. In response, Brame et al. (2012) argued that group-based analyses of developmental trajectories are useful for testing hypotheses based on theories about the existence of groups and about their characteristics and attributes.

In their recent discussion of group-based trajectory models for clinical research, Nagin & Odgers (2010a, p. 118) recommended eschewing the notion that latent classes represent “distinct entities” in favor of an understanding that trajectory groups are “just approximations of a more complex reality” and “clusters of individuals following similar trajectories.” Nevertheless, they proposed that such “clusters” are of scientific utility if they are “distinguishable in terms of pre-existing characteristics, subsequent outcomes, their response to treatment, or their relationship to trajectories for other outcomes or behaviors” (p. 118). In other words, groups that meet the above criteria are substantively important and perhaps worthy of their linguistic labels.

In *Psychological Methods*, Bauer & Curran (2003a,b; 2004) also questioned the utility of mixture modeling for life-course data. They relied on simulation studies to make two important observations. First, when the model assumptions are violated—e.g., when the actual outcome distribution is mildly nonnormal—a model with multiple trajectory groups could be consistently identified as optimal even though only one group was actually present (i.e., for a set of simulated data generated from a single underlying mean trajectory). Second, such overextraction of groups could have detrimental consequences for identifying the effects of predictors: Not only do tests have diminished power, but one could also obtain a spurious predictor because of class interaction, e.g., overextraction might falsely identify groups of individuals for whom an intervention differs in effectiveness. Bauer & Curran (2003a, p. 358) concluded with the following analogy: “The fact that multiple classes are optimal for the data no more indicates that the population is heterogeneous than a significant correlation indicates that Variable A causes Variable B.”

They further emphasized that, as for the correlation coefficient, the possibility for potentially incorrect interpretations does not necessarily diminish the utility of a method. However, although most researchers understand and appreciate the danger of misusing the correlation coefficient when making causal inferences, they do not have a similar understanding regarding the application of mixture models to life-course data. Bauer & Curran (2004, pp. 25–26) provided four recommendations: (a) determine the intended function of the latent classes, (b) attempt to ascertain the actual function of the latent classes, (c) embark on a program of construct validation, and (d) be careful when drawing conclusions from the fit of a single model. Bauer & Curran (2004) frame these insightful recommendations as open questions that go beyond a bulleted list of guidelines for practice to lay out a broad multiyear research agenda.

More recently, Hoeksma & Kelderman (2006) published commentary on a tutorial-style paper in *Infant and Child Development* by Connell & Frye (2006) that again raised the issue of identifying subgroups in a sample owing to violation of model assumptions and cautioned potential users that GMMs “do not offer a free way to growth patterns in unknown subpopulations” and should therefore be used only “with an understanding about the assumptions they make” (Hoeksma & Kelderman 2006, p. 632).

Finally, in the printed version of the 2004 Cattell Award Address, Bauer (2007, p. 759) described his own developmental trajectory with respect to GMM for life-course data. His careful analyses of the costs and benefits associated with using the methodology led him from exuberant excitement to pessimism. His conclusions (p. 782) were as follows:

The fundamental question I sought to address is whether these models are likely to advance psychological science. My firm conviction is that, if these models continue to be applied as they have been so far, the answer is clearly no. [...] I therefore believe that direct applications of GMMs should be refrained from unless both the theory and data behind the analysis are uncommonly strong. Otherwise, the application of GMMs in psychological research is likely to lead to more blind alleys than ways forward.

An alternative place for GMMs in psychological research is as an approximating device in indirect applications. [...] An interesting question, however, is whether GMMs will actually be used for this purpose when the idea of “real” groups seems to be so much more alluring.

To answer the question of whether applied research uses latent trajectory groups as an approximating device, Sterba et al. (2012) reviewed 100 randomly selected applications in psychology. They observed that although researchers typically mention that latent groups can serve as a semiparametric approximation of the underlying continuum, most applications do not consider whether such approximation is adequate or even necessary for their data.

5. LIFE-COURSE DATA ANALYSIS IN CRIMINOLOGY

Figure 1 illustrates the continued explosion of the number of applications of mixture models to life-course data despite the concerns discussed above. In this section, we provide a brief overview of criminological theories and debates about the relationship between age and crime, and we review two decades of applications of mixture models in criminology.

5.1. Theoretical Underpinnings

In the past 25 years, three theoretical perspectives have emerged in criminology concerning the shape and invariance of the age-crime curve, the existence of distinct groups of offending trajectories derived from criminal career or developmental theories, and the role of the life course in

shaping offending patterns. Criminological interest in trajectory analysis derives from the criminal career paradigm, which in turn has its roots in the Wolfgang et al. (1972) finding that 6% of a birth cohort of Philadelphia boys accounted for over 50% of the arrests in that entire cohort. This finding stimulated interest in the policy of selective incapacitation: If these 6% could be identified early in their careers, they could be separated from society, thereby reducing the crime rate in a cost-effective way. Blumstein & Cohen (1979) formalized this finding into a criminal career model by decomposing the crime rate into two components, participation (prevalence in the population) and frequency (incidence among active offenders). Subsequently, Blumstein and his colleagues tried to distinguish high-rate offenders from innocents and desisters, and they used Poisson models to estimate frequency rates among high-rate offenders to determine whether the offense rates persisted through the life span (e.g., Blumstein et al. 1985; Barnett et al. 1987). Later, Moffitt (1993) extended the distinction between high-rate offenders and desisters, renaming them life-course persistent and adolescence-limited offenders, respectively. She posited that life-course persistent offenders constitute less than 10% of the population, suffer from neurological deficits, and are at risk of antisocial behavior throughout the life span, whereas adolescence-limited offenders mimic the behavior of others as a normal part of adolescence and age out of crime as they take on adult roles.

Hirschi & Gottfredson (1983) challenged criminal career and typological approaches by arguing that the age-crime curve is invariant across all social groups and historical periods and is therefore beyond explanation by social theories. They claimed that crime results from stable individual differences in criminal propensities in conjunction with an exogenous effect of age and opportunity. Given these claims, they maintained that (a) research on typologies of trajectories is unnecessary because all groups follow the age-crime curve and (b) research on criminal careers is unnecessary because individual-level differences in the features of criminal careers, such as age at onset, age at peak offending, and age at desistance, can be explained by differences in criminal propensity. For example, they argued that apparent racial differences in age at onset merely result from higher criminal propensities among blacks than among whites. Furthermore, they concluded that longitudinal data are unnecessary for studying crime and that life-course transitions exert no discernible effect on criminal behavior.

Finally, life-course criminologists have emphasized the importance of life-course transitions in shaping offending trajectories. Sampson & Laub (1993), who were pioneers of life-course criminology, acknowledged the importance of individual differences in criminal propensities, which account for the relative stability of offending. However, they also argued that changes in offending over time result not only from changes in opportunities and age but also from life-course transitions that increase social bonds to conventional society. For example, it is not school, work, or marriage per se, but rather commitment to school, stable work careers, and quality marriages that produce desistance from crime. Furthermore, Laub & Sampson (2001) rejected the typological approach to modeling offending trajectories, arguing that empirically-derived offending groups fail to capture important theoretical differences, and that such groups are better conceptualized as discrete approximations of continuous individual differences in criminal propensity.

These three perspectives have distinct implications for child behavior problems from kindergarten to high school and beyond, such as those measured in the Montreal Longitudinal and Experimental Study. Typological perspectives predict the emergence of distinct groups of chronic versus adolescence-limited offenders, which would begin during the transition to high school. Life-course perspectives imply that trajectories may begin to diverge during school as some youths drop out, others perform poorly, and still others excel and become committed to education. Finally, perspectives emphasizing the invariance of the age-crime curve imply that all youth follow a single age-crime curve that differs only in terms of individual differences in stable criminal propensities

and time-varying criminal opportunities. Recent research using mixture models for life-course data has shed some light on these perspectives.

5.2. Review of Applications

Piquero (2008, p. 53) reviewed findings from 87 studies that used a mixture modeling approach to document the longitudinal patterning of criminal activity, noting that “for the most part, the trajectory methodology has been used to identify distinct trajectory systems and determine how they vary along relevant covariates”. After reviewing the main theoretical arguments and debates in criminology, Piquero (2008, pp. 49–50) summarized key substantive findings by the type of sample (offender-based or general population) and by the life-course time period (childhood, adolescence, or adulthood). He made two important observations:

1. Most studies identify between three and five groups that follow typical age patterns: “There tends to be a low rate group, a high rate group, a moderate but declining group, and a late onset group.”
2. The number of latent groups seems to be affected by the number of time points (more points result in more groups), spacing of observations (more frequent observations result in more groups), reporting method (self-reports result in more groups than do official records), outcome variable (delinquency scales generate more groups than do, e.g., police contacts), age-span in the study (longer spans result in more groups), and natural data sparseness (more frequent events, such as delinquency self-reports, generate more groups than do less frequent events, such as arrests).

Piquero’s first observation is supported by a more recent review of Jennings & Reingle (2012), who considered findings from 105 applications of mixture models to life-course data on violence, aggression, and delinquency. Jennings & Reingle (2012) observed that despite substantial variability in the number and shape of trajectory groups across samples, most studies identify a group of chronic offenders, a group of escalators or desisters, and a group of nonoffenders. Thus, the predominant differences among latent trajectory groups identified in criminology and developmental psychology appear to be in two key dimensions of variability in the group trajectories: the rate of offending and the timing and/or duration of the peak age of offending.

Piquero (2008) also observed that models tend to recover more groups when the data have larger variability, e.g., owing to measurement error (as in the case of self-reports), collection over a longer time span, more frequent measurement schedules, or other design variables affecting the number of available observations. Thus, researchers seeking to identify distinct subpopulations should be aware of the degree to which their results are sensitive to differences in study design, as cautioned by Bauer & Curran (2003a) and Eggleston et al. (2004). In the Section 8, we take this notion further and suggest that recovery of truly distinct subpopulations should be robust to minor changes in study design and that the burden of demonstrating this robustness should be on the researcher.

6. REVIEW OF STATISTICAL PRACTICE IN CRIMINOLOGY AND BEYOND

A distinctive characteristic of statistics, and a reason why “its tentacles reach into every corner of scientific enterprise,” is that it deals with variability and uncertainty (Bartholomew 1995, p. 5). We see life-course data analysis as no exception. It is crucial to communicate all sources of variability

in a model, including unexplained variability. The number and shapes of latent group trajectories allow researchers to describe only a portion of the variability in their data.

We conducted our own review of common statistical practices among applied studies using mixture modeling for life-course data, focusing on assumptions and findings about the type and amount of variability explained, the choice of the number of groups, and the shapes of group trajectories. Owing to the increasingly large volume of applications using mixture modeling for life-course data, we set out to thoroughly examine the 87 studies reviewed in Piquero (2008) and to supplement them with more recent papers. Thus, we selected a systematic sample of 20 papers from more than 200 published in 2007 and later. We made sure that our sample included both GMM and group-based trajectory applications published in criminology and in other areas and that at least 10 of the papers were published after 2012. In this section, we summarize our observations.

6.1. Variability

Debates about the merit of mixture modeling for life-course data center on questions about model assumptions and the structure of variability in individual trajectories. Therefore, we were particularly interested in examining whether and how these questions are being addressed in applied research. Our investigation did not focus on variations in technical specifications for modeling variability (e.g., GMMs versus group-based models, parametric versus nonparametric random effects); these choices are typically constrained by the technical capabilities of the software used for analysis. Instead, we focused on more fundamental issues of statistical practice: Were the assumptions about the outcome distribution and about the error terms clearly stated? Were error variance estimates, including residual variability, provided, and were the relevant implications of these estimates discussed? Were graphical methods employed to demonstrate different sources of variability (e.g., within-group and between-group variability)?

Given that individual heterogeneity has been a key motivation for developing mixture models, we were surprised to find that little attention has been paid to examining either observed or predicted variability. Approximately 20% of the 87 studies reviewed by Piquero (2008) did not clearly state their assumptions for the outcome distribution. More, of the ~30 papers that employed normal or censored normal outcome distributions, only Muthén & Shedden (1999) reported estimates for the error variability, σ_e . Many studies also omitted estimates for the variance-covariance matrix of random effects $\Sigma_v(k)$. The assumptions about variability of the random effects and residual error terms were rarely discussed, but some studies mentioned them when, for example, more general specifications of the error patterns appeared to interfere with convergence (e.g., Colder et al. 2001). In addition, there seems to be a common misconception that group-based trajectory models do not imply any variability around the group-specific means. However, this statement is incorrect because the existing software assumes residual errors for every model fit. These errors may be subsumed, as in the generalized linear model framework (e.g., for Poisson or negative binomial outcomes), or they may be specified explicitly (e.g., for normal or censored normal outcomes).

Not one paper among the early studies reviewed by Piquero (2008) used a graphical approach to illustrate variability of the observed and/or predicted individual outcomes around the mean trajectories of the groups. Instead, researchers typically provide graphs of observed and fitted means by group, similar to **Figure 3a**. This graph provides polynomial group mean trajectories estimated by a group-based trajectory model. [To be consistent with Nagin & Tremblay (1999) and Jones et al. (2001), who previously analyzed the Montreal opposition data, we used a censored normal distribution for the outcome.] This type of graph completely masks the variability of individual observations around the group means, creating an impression of distinct patterns. In contrast, Muthén et al. (2002) and Muthén (2004) use graphical visualizations that allow readers

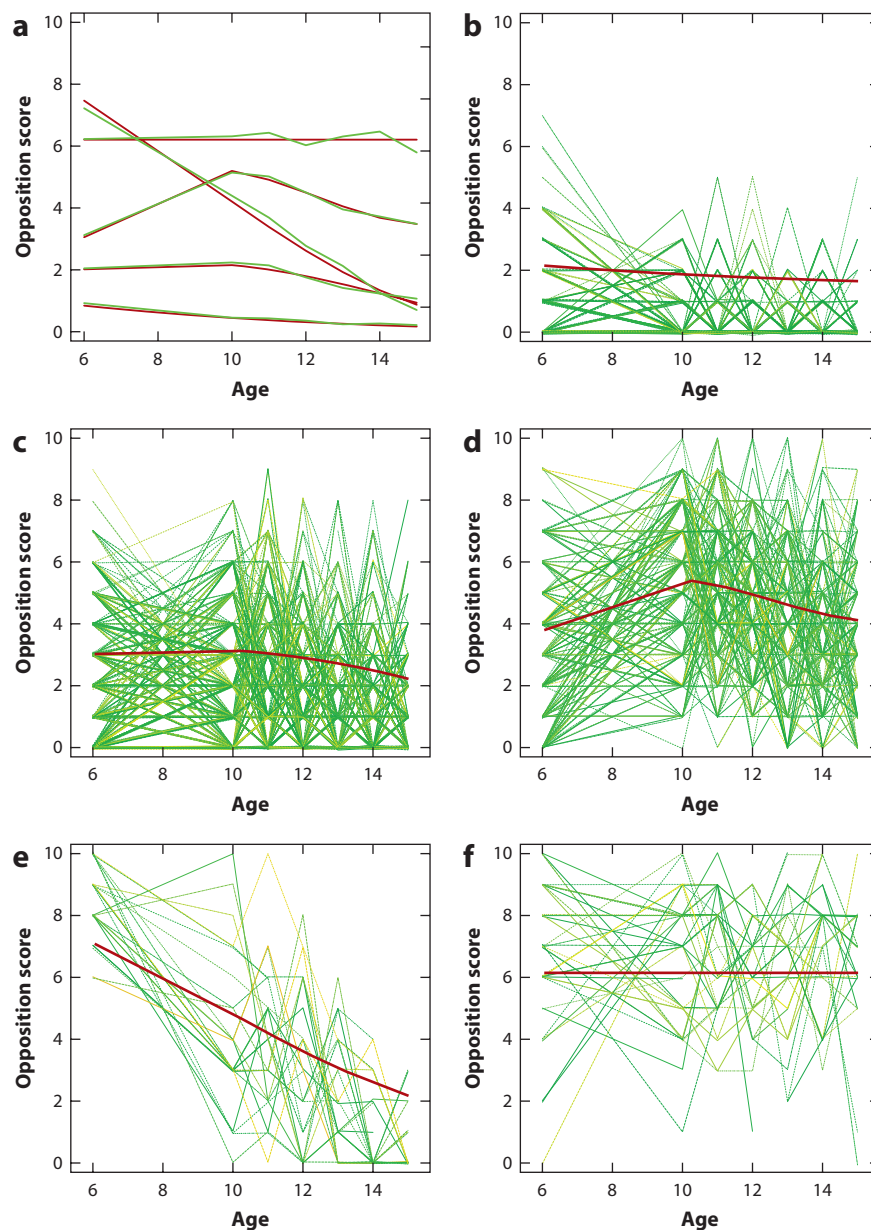


Figure 3

Latent class growth trajectories for the Montreal opposition data. A five-group model was identified as the best fitting model using the Bayesian information criterion. (a) Commonly used graphics showing predicted (red) and observed (green) group means. (b–f) Our recommended graphics, which show predicted group means (red) and observed individual trajectories (green) classified by trajectory shape for each of the five groups in panel a. Individual trajectories have been color-coded such that more intense colors correspond to higher posterior probabilities. To avoid superposing individual trajectories with identical responses, observed opposition scores that were integers were jittered by adding random noise.

to see the degree to which observed individual outcomes align with group mean trajectories. In addition, to examine model-based predictions, one might consider plotting confidence bands for individual observations around the group means (e.g., Bandeen-Roche et al. 1997).

As an illustrative example for continuous outcomes, consider the graphics for the Montreal opposition data in **Figure 3b–f**. Here, we have classified individuals into groups identified by the group-based trajectory model using the maximum posterior probability, but using random- or pseudo-class assignments can also be instructive (Bandeen-Roche et al. 1997). We then plotted predicted group mean trajectories along with the observed trajectories for the individuals classified into each group. These graphics allow us to visualize the extent to which individual variability is explained by group means, as well as the extent of overlap between observations of individuals from different groups.

The five-group model fits the data very well, according to guidelines described by Nagin (2005). For example, the probability that the five-group model is correct given the set of models with up to six groups is 0.97 (Kass & Wasserman 1995), and the posterior classification probabilities for most individuals are close to 1 (**Figure 3b–f**, *bright green lines*). Nonetheless, our visual examination indicates that within-group variability is high compared with variability among group mean trajectories. **Figure 3** indicates a lack of clear separation between individual trajectory observations from different groups: Alterations in an individual's trajectory that are minor in relation to within-individual variability could change his or her group membership. Hence, for the Montreal data, we conclude that the groups appear to serve as an approximation to an underlying continuum of variability in opposition trajectories. In this case, we cannot know whether the data represent multiple true groups with different overlapping distributions or a single group with a continuous distribution. **Figure 3a** shows that the trajectories of support for the underlying continuum differ primarily in the level of opposition and the peak age of opposition, but the peaks are not very pronounced for most groups.

We examined graphical illustrations in more than 200 papers on mixture models for life-course data published in 2007 or later, only nine of which graphically depicted individual trajectories within groups: Boscardin et al. (2008), Bowers & Sprott (2012), Leoutsakos et al. (2012), Odgers et al. (2008), Qureshi & Fang (2011), Reinecke & Seddig (2011), Rolfe et al. (2011), Saunders (2010), and Underwood et al. (2009). Many of these publications used the graphics as some sort of diagnostic tool. For example, Underwood et al. (2009) examined individual trajectories and the overall mean curve, observed high individual heterogeneity, and concluded that a mixture model was needed to describe their data. Similarly, Qureshi & Fang (2011) examined individual trajectories around group means to determine whether to move from fitting group-based trajectory models to fitting GMMs. Saunders (2010) employed an analogous type of plot to examine individuals whose class assignments “shifted” between two different versions of a model. Boscardin et al. (2008) used observed individual trajectories around group means to examine the overall fit quality, and Odgers et al. (2008) plotted the predicted individual trajectories by group to show that patterns among males and females are similar.

In all but one of the nine publications discussed above, individual variability around group means was high compared with the observed differences among group mean trajectories. This evidence suggests that individual trajectories are distributed along an underlying continuum, as opposed to among distinct and well-separated groups. Surprisingly, none of the authors use their graphics to make this point, and only one paper acknowledges the need to view latent groups as trajectories of support for the underlying latent continuum (Odgers et al. 2008).

We found that only the study by Qureshi & Fang (2011) on socialization trajectories in open-source software projects demonstrated truly distinct patterns of development among the groups obtained from the best-fitting model. A visual examination of figure 4A in their article shows that

the within-group variability is less than the between-group variability. In this case, the groups are well separated in time: There exists a time interval during which individual trajectories from different groups do not overlap. However, the separation between fitted groups may be manifested in a more general sense in the multivariate parameter space of a model; separation between groups does not necessarily have to occur in time. In fact, the groups can still be thought of as distinct despite substantial overlap among observations from different groups at each point in time. One such case is that in which one group shows a steady decline and has a relatively small variability in individual trajectories around the group mean compared with no change in the other groups. We return to the identification of distinct groups in Section 8.

Interestingly, when Qureshi & Fang (2011) found four truly distinct groups of socialization trajectories, they also observed that in fitting GMMs with more than four groups, the individuals were still placed into four classes (i.e., additional fitted classes were empty) in the cases of successful convergence. This observation may be related to recently proven asymptotic properties of simpler mixtures of normals. Thus, Rousseau & Mengersen (2011) considered the case in which the number of fitted components in the mixture exceeds the true number of components and proved that the posterior distribution in this case has a stable behavior, concentrating on the true components and having zero weights for the extra components.

6.2. Choice of the Number of Groups

Almost every review or tutorial paper we examined touched on the question of choosing the dimensionality of mixture models for life-course data. Nagin (2005) provided a number of recommendations for choosing an appropriate dimensionality, including penalized likelihood criteria such as the Akaike information criterion (AIC) (Akaike 1973) and the BIC (Schwarz 1978), as well as recommendations to focus on groups with distinct trajectories and to consider substantive knowledge. The statistical literature on finite mixture models has criticized the AIC and BIC as inapplicable for assessing the number of mixture components owing to violations of the regularity conditions (McLachlan & Peel 2000). Instead, Muthén (2004) and Muthén & Asparouhov (2008) suggest using a bootstrapped likelihood ratio test (see McLachlan & Peel 2000), among other alternatives.

Most published studies fitted models with different numbers of groups, K , and then decided on dimensionality (i.e., the optimal number of groups). However, the methods for choosing the upper bound for K , as well as for choosing the optimal number of groups, are often unclear. Several studies, predominantly those that employed group-based trajectory modeling, observed that the goodness-of-fit criteria kept improving as the number of groups increased until convergence problems arose (e.g., Brame et al. 2001; D'Unger et al. 2002; Li et al. 2001; Piquero et al. 2001, 2002; Wiesner & Capaldi 2003). Several studies declared the optimal number of groups to be the largest K from a given set but did not comment on analyses with higher numbers of groups (e.g., Hamil-Luker et al. 2004, Nagin & Tremblay 1999, Paternoster et al. 2001). When the optimal dimensionality identified by goodness-of-fit criteria is perceived as large, applied researchers tend to reduce it to a lower number of groups that would still represent what they believe are substantively distinct trajectory patterns. In particular, researchers typically remove groups that appear to account for only minor variations in level or timing (e.g., Blokland et al. 2005, Bushway et al. 2003, Lacourse et al. 2002, Nagin & Tremblay 2001).

6.3. Trajectory Shapes

Each of the most prominent trajectory shapes identified in applied criminology research on life-course trajectories, as summarized by Piquero (2008), can be described as weakly unimodal—that

is, as a trajectory or part of a trajectory that has only one hump. Importantly, weakly unimodal trajectories include those that reach their peak outside of the observation interval, those that plateau for extended periods of time, and even those that exhibit little change over the entire life course. Weak unimodality is consistent with descriptive studies of various age-crime curves (e.g., Hirschi & Gottfredson 1983). In addition, after reviewing graphical representations of the mean group trajectories in the 87 early studies considered by Piquero (2008), we found that most group curves do appear to be weakly unimodal in shape. Studies that identified shapes that do not fit this description either used data that were not related to criminal or antisocial behavior problems (e.g., women's employment trajectories just before and after a baby's birth appear to be bimodal with prebirth and postbirth peaks; see Hynes & Clarkberg 2005) or obtained bimodal curves with secondary modes on the boundary (e.g., Bushway et al. 2003, Chassin et al. 2004, Wiesner & Windle 2004). These secondary modes could have been artifacts that resulted from a lack of data before or after the observed time interval. We should note that the observed unimodality in fitted group trajectories from growth curve and group-based trajectory analyses might be driven partly by the use of low-degree polynomials. This in turn may mean that a low signal-to-noise ratio, in conjunction with potential model misspecification issues, makes estimating more complex functional forms difficult.

6.4. Summary

Three general points emerge from our review of statistical practice. First, either observed or model-predicted within-group variability around mean trajectories is rarely examined. Moreover, the typical examples of graphics used in these analyses give little information about the sources of variability, including which sources are and are not explained by the models. Second, in choosing the number of groups, researchers often depart from the use of model fit indicators; they tend to stop considering additional groups when the addition of a group produces convergence problems or when additional groups appear to add little substantively. In the former case, lack of convergence at higher dimensions does not imply optimality at a lower dimension but may rather signal lack of consensus between the data and the model. In the latter case, additional groups are often discarded in practice because they differ only slightly in level or timing, possibly biasing the results in favor of a hypothesis that the groups are distinct over the alternative that the groups are, approximating a continuous distribution. Third, the estimated trajectory groups in criminology exhibit weakly unimodal shapes with some differences in the level and timing of offending. This finding is consistent with the age-invariance thesis proposed by Hirschi & Gottfredson (1983). Empirically identified trajectory groups do not reveal the life-course persistent group with a constant rate of offending that criminal career and dual taxonomy approaches predict.

7. UNIMODAL CURVE REGISTRATION FOR INDIVIDUAL TRAJECTORIES

In this section, we review an alternative approach to life-course data analysis that is based on explicit modeling of individual criminal activity levels and timing of criminal activity (see Telesca et al. 2012). The ability to model variability in individual time schedules sets this approach apart from growth curve models and trajectory mixture models, both of which lack an explicit way of specifying subject-specific variation in the timing of crime. Individual time schedules reflect individual variation in time-related features of criminal careers such as the age at first offense, the age at peak criminal activity, and the age at desistance from crime. This variability is explained by subject-specific behavioral clocks that can be related to each other to establish a typical timing of crime. We illustrate this concept in **Figure 4**.

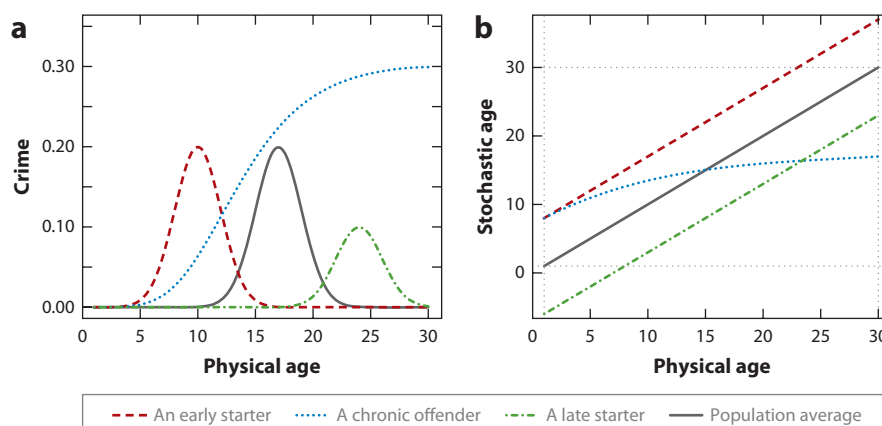


Figure 4

Unimodal curve registration example. (a) Individual crime trajectories. The solid dark gray curve defines a typical criminal career. The dotted red line depicts the trajectory of an individual whose criminal career follows the exact same shape as the mean trajectory but starts early, shifting the trajectory to the left. This trajectory corresponds to a behavioral hypothesis that some individuals start offending early but follow a typical age-crime curve. In contrast, the dotted blue line depicts the trajectory of an individual who starts early but whose criminal career also follows an altered shape: It reaches a peak late, and the individual maintains peak offending for a longer period of time than does an individual following a typical trajectory. (Individuals following this trajectory maintain peak offending until at least age 30.) This trajectory might correspond to the behavior of a chronic offender. Finally, the dotted green line depicts an individual whose offending trajectory starts late and who offends less than the population average. (b) Corresponding individual time transformation functions for the trajectories in panel a.

The lack of explicit mechanisms for modeling individual variability in time schedules may explain the empirical observations that growth curve and trajectory mixture models are typically unable to capture individual variability in the decline of offending (Bushway et al. 2009). From a statistical perspective, the model misspecification that occurs when individual variability in timing is present in the data but ignored at the modeling stage results in inconsistent estimates of time-varying population parameters and in population estimates of age-crime curves that are not representative of typical individual trajectories (Brumback & Lindstrom 2004, Gervini & Gasser 2004). An illustration of this phenomenon appears clearly in trajectories of yearly counts of marijuana smoking (Figure 5). The raw pointwise mean age-crime curve in Figure 5a smears over individual peaks of marijuana use, whereas the age-crime curve in Figure 5b, which was estimated from individual trajectories that were aligned to a common time schedule, has a more pronounced peak and decline of marijuana smoking activity. We discuss this example in more detail in Section 7.2. Because the estimation of time transformation functions allows individual trajectories to be aligned to a typical time course, the statistical literature traditionally refers to the technique as curve registration.

7.1. Curve Registration Models

Curve registration models have been used extensively in analyses of biomedical data to align sampled trajectories (Brumback & Lindstrom 2004, Gervini & Gasser 2004, Ramsay & Li 1998). Biomedical measurements exhibit natural variability in phase and amplitude, are typically sampled in large quantities on a fine grid in time or space, and are rarely subject to large

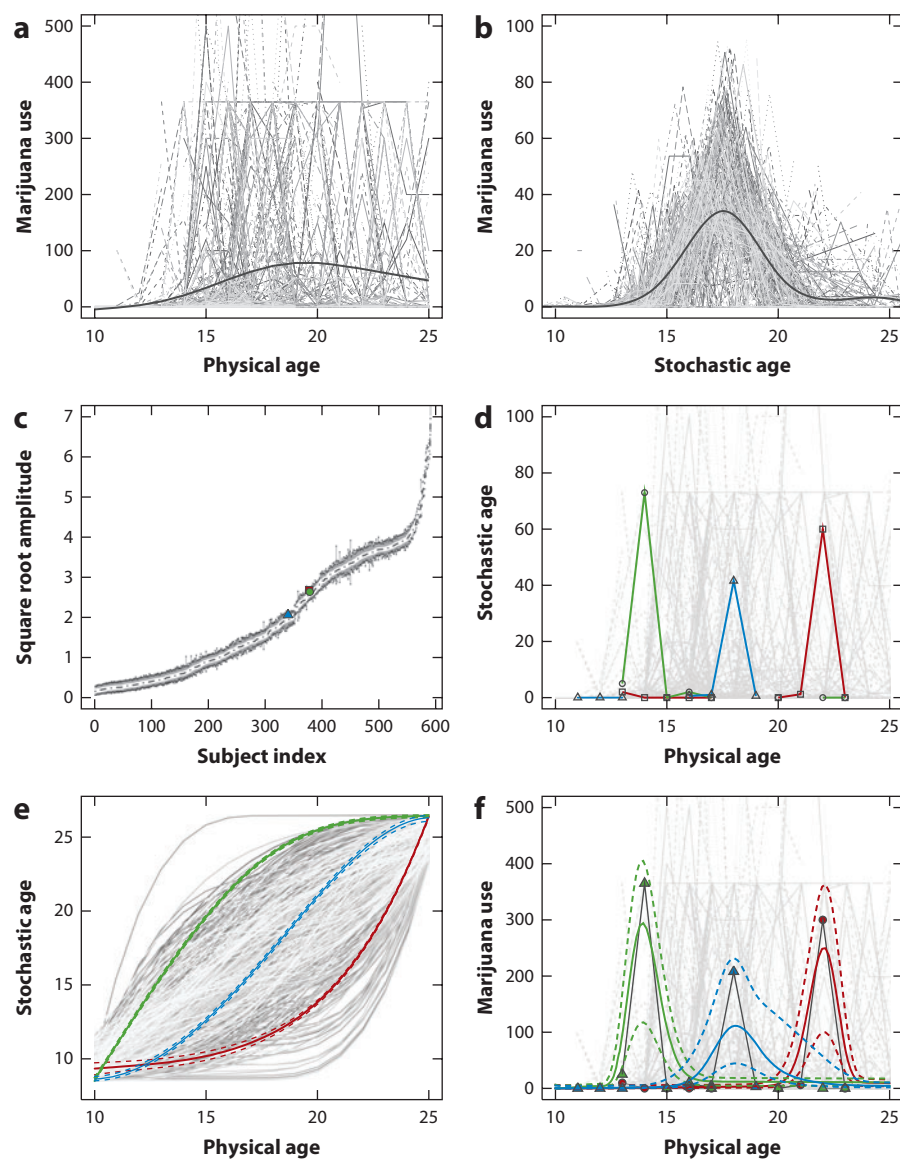


Figure 5

Marijuana use. (a) Individual trajectories and pointwise average age-crime curve. (b) Aligned normalized trajectories with a unimodal curve registration (UCR)-estimated age-crime curve. (c) Individual amplitude estimates and 95% credible intervals. Colored symbols represent amplitudes for the subjects highlighted in panel (d). (d) Observed yearly counts of marijuana smoking. Counts for three subjects exhibiting different time schedules are highlighted in color. (e) Posterior mean time transformation functions. Colored curves represent individual time schedules (solid lines) and 95% credible intervals (dashed lines) for the subjects in panel (d). (f) Observed (solid dark gray lines) and predicted (solid colored lines) yearly counts for subjects in panel (d), along with associated 95% prediction intervals (dashed colored lines). Parts of this figure have been adapted from Telesca et al. (2012), reprinted by permission of Taylor & Francis Ltd., www.tandfonline.com.

measurement errors. Telesca & Inoue (2008) first formulated a Bayesian hierarchical model for curve registration, which allows for the estimation of individual-specific time scales (warping functions). Later, Telesca et al. (2012) extended this idea to a statistical framework for the analysis of count data in criminology, including covariate effects on the expected intensity of criminal behavior and on the deviation from the average timing of offenses. They propose borrowing information across curves in a Bayesian framework and placing structural constraints on the shape of the age-crime curve to account for data sparsity and measurement errors that are common to life-course data on behavioral trajectories.

Assuming a unimodal population age-crime curve, Telesca et al. (2012) developed a Poisson warping regression model in which individual crime trajectories are random functions that deviate from the population age-crime curve according to individual-specific levels of offending and temporal misalignment. Unlike other models of life-course data, the UCR approach explicitly incorporates the criminological arguments that the population age-crime curve is unimodal and that individual time schedule variations contribute to the departures of their trajectories from the common population curve.

In this section, we provide two illustrations of the UCR approach for count and scale outcomes using yearly marijuana smoking counts from the Denver Youth Survey (DYS) (Esbensen & Huizinga 1990) and opposition scores from the Montreal Longitudinal and Experimental Study. For the latter, we used a censored normal for the scale outcome to be consistent with previous recommendations (e.g., Jones et al. 2001).

7.2. Poisson Unimodal Curve Registration

In the following subsections, we describe the UCR model for count data with covariate effects on the individual-specific amplitude and timing of criminal activity, as originally developed in Telesca et al. (2012). We then briefly review their original analysis of marijuana use data from the DYS.

7.2.1. Model description. Let $\mathbf{y}_i = (y_{i1}, \dots, y_{ij}, \dots, y_{im})'$ denote the outcome for subject i (where $i = 1, \dots, n$) over a discrete sampling design in time $\mathbf{t} = (t_1, \dots, t_j, \dots, t_m)'$. For count outcomes, y_{ij} is the number of offenses over a time interval τ , which is fixed and held constant for all observations and individuals (one year in our case). If \mathbf{X}_i denotes a p -dimensional vector of covariates associated with subject i , we let $E(y_{ij}) = \lambda_i(t_j, \mathbf{X}_i)$ be the expected offense level for subject i at time t_j . The following sampling model (Telesca et al. 2012) is assumed:

$$y_{ij} | \lambda_i(t_j, \mathbf{X}_i) \sim \text{Poisson}(\lambda_i(t_j, \mathbf{X}_i)). \quad 3.$$

The subject-specific mean $\lambda_i(\mathbf{t}, \mathbf{X}_i)$ is defined as a population age-crime curve $M(\mathbf{t})$ multiplied by a subject-specific intensity of offense $a_i(\mathbf{X}_i) \geq 0$ and evaluated over a subject-specific time scale $u_i(\mathbf{t}, \mathbf{X}_i)$. The model in Equation 3 therefore defines a Poisson warping regression with intensity function

$$\lambda_i(t_j, \mathbf{X}_i) = a_i(\mathbf{X}_i)(M \circ u_i)(t_j, \mathbf{X}_i) = a_i(\mathbf{X}_i)M\{u_i(t_j, \mathbf{X}_i)\}. \quad 4.$$

In the foregoing formulation, both the intensity and timing of criminal activity are made dependent on covariate information, assuming $E(a_i) = \exp\{\mathbf{X}_i' \mathbf{b}_a\}$ and $E(u_i(\mathbf{t})) = \mathbf{t} + \mathbf{X}_i' \mathbf{b}_u$. The model is naturally described in terms of a three-stage hierarchical model for which Bayesian estimates of both subject-specific and population-level quantities are easily obtained (for details, see Telesca et al. 2012).

7.2.2. Example: marijuana smoking data from the Denver Youth Survey. The DYS is a study of delinquency and drug use in high-risk neighborhoods in Denver. The DYS collected data via an accelerated longitudinal design covering the age span from 7 to 25 years old. The survey asked questions about drug use starting from age 11. Subjects were interviewed in their homes annually from 1988–1992 and 1995–1999 (10 waves). Telesca et al. (2012) consider answers to the survey question: “In the past year, how many times have you smoked marijuana?”

Longitudinal observations of marijuana use are reported in **Figure 5a**. A few observations indicating that a subject had smoked marijuana more than 500 times per year have been cut off for ease of visualization. **Figures 5c,e** report the estimated amplitude parameters and warping functions, respectively, for all participants. **Figure 5d** highlights individual observations for three marijuana smokers: an early user, an average user, and a late user. **Figure 5e** shows the posterior mean time transformation functions and associated 95% credible bands for each participant, and **Figure 5f** shows their predicted trajectories and associated 95% prediction intervals. For the average user, the time transformation is close to identity, i.e., the stochastic age of this person is similar to his or her physical age. The late user’s stochastic age is roughly constant until he or she reaches a physical age of approximately 18; this individual then goes through the marijuana use period much faster than does an average marijuana smoker in our sample. Similar to the late user, the early user exhibits a fast period of marijuana smoking, but the period spans at much earlier ages.

The analysis conducted by Telesca et al. (2012) confirms that the frequency and timing of marijuana use are two large sources of variability, which they then model as a function of race and gender. They find that, as expected, females use marijuana less frequently. For example, white females smoke marijuana approximately 0.72 times as frequently as do their male counterparts. African-American males, however, appear to use marijuana more frequently (1.55 times as often) than their Caucasian counterparts. Similarly, the study reports significant differences in the timing of drug use; on average, females start using marijuana approximately 5.5 months earlier than do white males, and African Americans start using marijuana approximately 10 months later.

This modeling approach allows one to estimate a common age-crime curve and to fit individual trajectories as departures from that mean curve. Furthermore, it allows one to disentangle variation in individual crime trajectories that results from differences in the level (amplitude) and timing (phase shift) of offenses, thereby providing new insights into modeling trajectories of marijuana use. Similar to the results of previous research, Telesca et al. (2012) find race and gender differences in levels of marijuana smoking. However, in contrast to previous research, they find little support for the claim that differences in the shape of the age-crime curve are merely a result of differences in rates of offending (Hirschi & Gottfredson (1983)). By explicitly disentangling observed variability into separate sources, Telesca et al. (2012) provide an approach that allows researchers to conduct statistical tests on important quantities of interest—tests that were not possible given traditional modeling of longitudinal trajectories. Thus, they find significant race and gender differences in timing (while controlling for comparable differences in amplitude), as indicated by the shift of marijuana use trajectories.

The assumption of unimodality may not be warranted in all applications. Telesca et al. (2012) compared the predictive performance of the UCR model against that of a curve registration model that uses a shape-unrestricted age-crime curve via a posterior predictive loss criterion, and examined the potential lack of fit for individual trajectories by computing conditional predictive ordinates (Pettit 1990, Geisser 1993). They find that, for the DYS marijuana smoking data, the predictive performance of the UCR model is superior to that of unconstrained curve registration models.

7.3. Censored Normal Registration

In this section, we extend Telesca et al.'s (2012) UCR model to the case of censored normal data and apply it to an analysis of the Montreal opposition data. As in Section 7.2.1, we model expected individual opposition trajectories as random amplitude and timing departures from a common unimodal age-opposition curve. The exploratory summaries in **Figure 2** highlight significant elements of individual heterogeneity: variation in opposition levels (**Figure 2a,b**) and timing (**Figure 2c**). In contrast to the formulation in Section 7.2.1, however, this analysis does not include covariate information.

7.3.1. Model description. Let y_{ij} denote the observed opposition score for subject i at time t_j . We follow Nagin & Tremblay (1999) and consider a censored normal sampling model for which the left censoring point η_0 and right censoring point η_1 are common to all individuals. Let $\mu_i(t_j)$ denote the mean parameter evaluated at time t_j for subject i , and let σ_ϵ^2 be the variance. Furthermore, let y_{ij}^* denote an uncensored latent variable (Chib 1992). The opposition score at time t_j for individual i is modeled as follows:

$$y_{ij}|y_{ij}^*; \eta_0, \eta_1 = \min[\max\{\eta_0, y_{ij}^*\}, \eta_1]; \quad y_{ij}^*|\mu_i(t_j), \sigma_\epsilon^2 \sim N(\mu_i(t_j), \sigma_\epsilon^2). \quad 5.$$

We assume that the mean structure $\mu_i(t_j)$ is defined in terms of a random affine transformation of a population mean function $M(t_j)$ evaluated over individual-specific time scales $u_i(t_j)$. More precisely, let c_i denote random intercepts, and let $a_i \geq 0$ denote subject-specific amplitude parameters. The time transformation functions $u_i(t) : [t_1, t_m] \rightarrow [t_1 - \delta, t_m + \delta]$ are assumed to define a bijection between the observed time scale \mathbf{t} and a subject-specific schedule $u_i(\mathbf{t})$. Furthermore, the population mean over time is defined in terms of a random function $M(t) : [t_1 - \delta, t_m + \delta] \rightarrow \mathbb{R}$. We assume that the mean trajectory for individual i follows the functional form

$$\mu_i(t_j) = c_i + a_i(M \circ u_i)(t_j). \quad 5.$$

The effect of individual time transformation $u_i(\cdot)$ on $M(\mathbf{t})$ is illustrated in **Figure 6**, in which we highlight the trajectories of two subjects (**Figure 6d**) and their associated time schedules (**Figure 6e**).

We represent the random functions $M(\cdot)$ and $u_i(\cdot)$ as linear combinations of B-spline basis functions (De Boor 1978). In the Bayesian setting, it is natural to represent this model in a hierarchical fashion (Telesca et al. 2012). Posterior computation is based on Markov chain Monte Carlo simulation achieved by combining the Gibbs algorithm presented in Chib (1992) with the Gibbs–Metropolis–Hastings sampler described in Telesca & Inoue (2008).

7.3.2. Example: opposition score data from the Montreal Longitudinal and Experimental Study. We restricted our analysis to subjects for whom the data included at least five, not necessarily consecutive, observations during the course of the study.¹ After removing 57 subjects with fewer than five observed time points, we were left with a subset of 859 individuals. Our inferences are based on 15,000 (thinned by 20) samples from the posterior distribution, after discarding a conservative 50,000 iterations for burn-in. Results are very robust to the choice of priors and number of basis functions.

¹The inclusion of observations with a shorter time series would not affect the population estimates; however, posterior inferences on subjects with fewer than five records could be misleading owing to weak identifiability of the subject-level parameters.

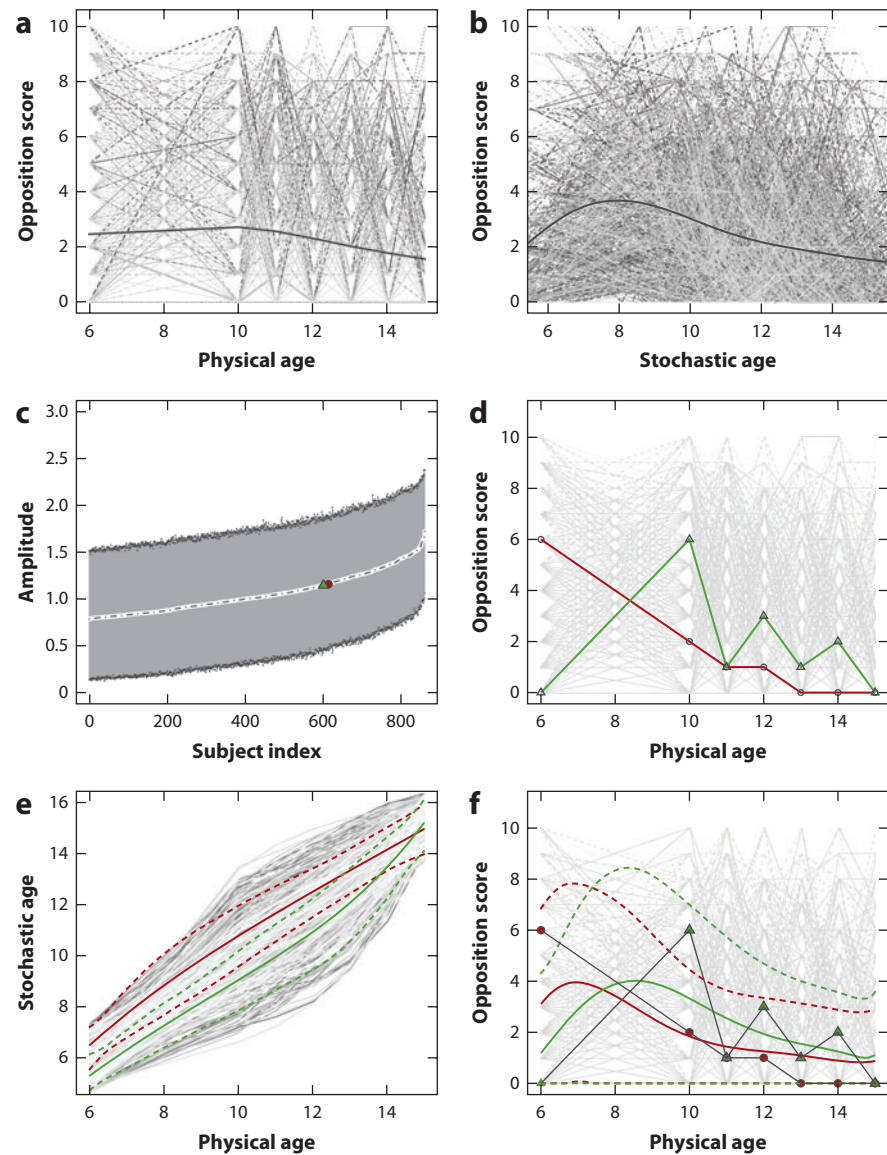


Figure 6

Opposition data. (a) Individual trajectories and pointwise average age-crime curve. (b) Aligned normalized trajectories with a unimodal curve registration-estimated age-crime curve. (c) Individual amplitude estimates and 95% credible intervals. Colored symbols represent amplitudes for the subjects highlighted in panel (d). (d) Observed opposition scores. Colored lines indicate scores for two subjects exhibiting different time courses. (e) Posterior mean time transformation functions. Colored lines indicate individual time schedules (solid lines) and 95% credible intervals (dashed lines) for the subjects highlighted in panel (d). (f) Observed (solid dark gray lines) and predicted (solid colored lines) opposition scores for subjects from panel (d), along with associated 95% prediction intervals (dashed colored lines).

We summarize results of our analysis graphically in **Figure 6**. **Figure 6a** provides the raw observed opposition scores and the pointwise smoothed mean. The raw opposition scores show high variability, and the overall pointwise smoothed mean curve does not have a pronounced peak. In contrast, **Figure 6b** shows that observed opposition scores aligned to the common time schedule resemble a haystack, and the UCR-estimated population age-crime curve peaks at approximately eight years of age, followed by a slow, constant decrease in the population score until the early teenage years. Estimated individual amplitudes have large uncertainties (**Figure 6c**), as do individual time-transformation schedules (**Figure 6e**); we invite the reader to compare the estimated uncertainty estimates in **Figure 6c,e** with those for the marijuana smoking data in **Figure 5c,e**, respectively. The predicted uncertainty for the observed trajectories of the two individuals highlighted in **Figure 6f** is also substantial (compare with **Figure 5f**): The 95% posterior confidence bands for the two individuals cover a large percentage of all observed data points. We checked the adequacy of the model by comparing the posterior predictive mean and maximal opposition levels with their observed counterparts (not shown), and we noted that the model recovers individual values well; individuals with maximum opposition scores of zero may be an exception because shrinkage toward the population mean predicts that even these subjects would exhibit positive scores. We found that estimates of shape-unrestricted age-opposition curves are highly sensitive to the priors and to the B-spline specifications, indicating potential identifiability issues. In contrast, unimodality restrictions provided us with a robust analytical framework without compromising predictive accuracy.

Overall, the high level of uncertainty in the model predictions illustrated in **Figure 6** indicates a low signal-to-noise ratio for estimating the individual level and timing of opposition trajectories. A similar conclusion could be reached by examining individual variability around the group means from the group-based model in **Figure 3a–f**. We ask the reader to compare this low signal-to-noise finding with the conclusion of Jones et al. (2001, p. 380), who interpret figure 2 of their article (which is analogous to **Figure 3a** here) by providing estimated percentages of subjects who exhibit “little or no oppositional behavior,” “low and somewhat decreasing levels of oppositional behavior,” “moderate levels of oppositional behavior,” “chronic problems of oppositional behavior,” and “high levels of oppositional behavior that drop steadily with age.” High within-group variability relative to between-group variability indicates that the number, shapes, and percentages of the groups identified with the Montreal data are not likely to be robust to even minor changes in study design (i.e., replicating the number of groups, their trajectory shapes, and their respective population percentages may be difficult).

8. DISCUSSION

8.1. Overview

In the preceding sections, we focused on finite mixture models for life-course data, discussed their application within criminology and developmental psychology, reviewed recent controversies over their substantive use and interpretation, and summarized important issues of statistical practice and the challenges these models raise. We also presented the most recent method of examining behavioral trajectories in criminology, the UCR approach, and we briefly compared it to finite mixture models using a classic data set from developmental psychology.

As our review emphasizes, different models for analyzing individual trajectories of behavior problems have distinct strengths and weaknesses for addressing different substantive questions. Generalized linear mixed models can be useful for examining a continuous distribution of trajectories but may have a limited ability to describe complex distributions. Group-based trajectory

models allow for mixtures of groups with different trajectory means (Nagin 2005). Generalized GMMs are similar to generalized linear mixed models, but they allow one to account for within-group heterogeneity using flexible covariance structures for within-group random effects (Muthén & Asparouhov 2008).

Whether they rely on a single population mean or group-specific means, applications of these three types of models to life-course data typically assume that individual trajectories are polynomial in age and may have polynomial random effects such as age and age squared, although other functional and semiparametric specifications can be used. Independently of the mean structure specification, these models do not have explicit mechanisms to capture heterogeneity due to variability in individual time schedules, termed phase variability. Failure to account for phase variability in the data can have enormous substantive implications because time-varying population estimates will be inconsistent unless the individual trajectories have been properly aligned to a standardized schedule (Brumback & Lindstrom 2004, Gervini & Gasser 2004, Telesca & Inoue 2008). Thus, without adjusting for phase variability, estimates of individual quantities related to timing, such as age at peak use, may not be reliable, and findings of multiple latent groups of individual trajectories may be spurious.

In contrast, the UCR model formulation *explicitly* accounts for variability in individual amplitude (level of offending) and timing and can include covariate effects on these quantities (Telesca et al. 2012). The model draws on theoretical arguments in criminology (Hirschi & Gottfredson 1983) by specifying a population-invariant age-crime curve for which individual departures are based on differences in crime propensity and life situations. These differences are reflected in individual-specific offending levels and time schedules, respectively. Although the assumption of unimodality may not be applicable in other contexts, e.g., when studying progression of a disease, the functional constraint improved the robustness of Bayesian curve registration approaches in our applications for marijuana use and opposition score data.

We are not suggesting that our UCR model of behavioral trajectories is superior to other trajectory models or that it should be treated as the method of choice. Instead, we are simply proposing an alternative approach that has distinct strengths and weaknesses relative to other methods. Perhaps the most obvious strength is that it focuses attention on an empirical pattern revealed by group-based models and allows for individual differences in time schedules and offending levels. However, similar to hierarchical linear models, the UCR approach has the potential weakness that shrinkage of the individual estimates toward the time-aligned population mean may make it difficult to recognize distinct clusters. Thus, a promising direction for future research would be to investigate whether functional data can be fruitfully clustered into groups after accounting for individual variability in offending level and timing. A simple approach to this question would be to use a clustering technique on the coefficients obtained from a UCR model, similar to a method suggested by Tarpey (2007) for functional data. Even better would be a more complex hierarchical approach that can combine UCR and clustering models.

8.2. On the Use of Mixture Models for Life-Course Data

As noted above, applications of mixture models for life-course data have exploded in the literature, to the degree that some argue that these models have become the statistical method of choice for criminologists and developmental psychologists. We believe that mixture models for life-course data represent an important methodological advance for analyzing behavioral trajectories. By incorporating finite mixtures into the model, group-based methods can, in principle, provide important information about the distribution of individual trajectories. Mixture

models have provided new ways of thinking about individual heterogeneity in longitudinal data and, consequently, new ways of conceptualizing substantive issues.

In their original introduction of a group-based finite-mixture trajectory model, Nagin & Land (1993) noted that the latent classes in group-based trajectory models can be conceptualized in one of two ways: as approximations of a continuous, unknown, and possibly complicated distribution of population heterogeneity or as concrete groups that can be treated as distinct and substantively useful entities. This distinction has far-reaching implications. Identification of subpopulation groups lends legitimacy to them, posing the danger that they will be viewed as more than mere approximations or heuristic devices (Sampson & Laub 2003). Moreover, as Nagin & Tremblay (2005a, p. 883) warn, group reification may be most acute in policy circles:

If a group is small and its behavior is socially undesirable, such as committing crimes, reifying the group as a distinct entity—rather than as an extreme on a continuum, may provoke draconian responses to the behavior by creating the impression of a bright line of separation between “them” and “us.” Human history is replete with tragic instances in which a fictional group-based separation is the first step in dehumanizing them.

Thus, especially when latent groups cannot be identified reliably, treating them as distinct entities with labels such as “superpredators” and “chronic offenders” is more than an issue of semantics. For example, in criminology, the concept of a chronic or high-rate offender has fueled interest in policies of selective incapacitation (Blumstein & Moitra 1980, Greenwood & Turner 1987).

Although latent trajectory groups are first and foremost an approximation to an unknown distribution of population trajectories, this view is underutilized. As approximations, latent groups of trajectories can be used in two ways: to provide valuable descriptive information on the distribution of trajectories in the population and to control for unobserved heterogeneity. Sterba et al. (2012) provided one example of the former in an article that focused on the use of latent groups to approximate the overall distribution of observations. Nagin & Land (1993) provided an example of the latter when they used latent groups to control for population heterogeneity in models of criminal careers. We have found few other examples that explicitly use mixture models for life-course data to control for unobserved heterogeneity in individual trajectories. The scarcity of such studies is unfortunate because mixture models for life-course data have the potential to address perhaps the most important specification problem in applying multivariate models to nonexperimental data on individual behavior: omitted variable bias or spuriousness due to unobserved confounders. As Heckman & Singer (1984) showed for hazard models, corrections for omitted variable bias that control for unobserved heterogeneity are highly sensitive to the distributional assumptions for the heterogeneity component. Using latent class models, Heckman & Singer (1984) provided a nonparametric approach to approximate unobserved heterogeneity. Nagin & Land (1993) and Land et al. (2001) showed that latent classes from mixture models for life-course data can be used in precisely the Heckman & Singer (1984) way.

However, our review indicates that the second conceptualization of latent groups—which allows researchers to treat these groups as distinct subpopulations—has been the prevalent one in the applied literature. We agree with Nagin (2005) and Nagin & Tremblay (2005a, p. 879) that distinct groups are rare in the social sciences:

Although there may be populations made up of groups that are literally distinct, they are not the norm. Most are composed of a collection of individual-level developmental trajectories continuously distributed across population members.

Nevertheless, most applications of group-based trajectory methods tend to reify the groups: Groups are given descriptive names, discussed as distinct entities, and considered in terms of policy implications. However, these studies rarely examine the extent of within-group variability in comparison with that of between-group variability and the residual error. Such empirical studies account for the overwhelming bulk of the growing number of applications of mixture models to life-course data. Given the current state of substantive theory about group heterogeneity in trajectories of crime and antisocial behavior, we suspect that discovery of distinct groups of trajectories will be relatively rare. With this caveat, the following section provides several recommendations for researchers interested in finding distinct groups of trajectories.

8.3. Recommendations

We believe that identification of distinct groups in life-course data requires careful examination of the extent of between- and within-group variability in the trajectories. Using the opposition score data, we showed that examining plots of individual trajectories by latent groups can shed important light on whether individuals from different trajectory groups follow distinct and well-separated patterns. Researchers who are interested in recovering true subpopulations must demonstrate that the group trajectories are distinct. Related to this, more methodological work needs to be done to allow researchers to delineate the conditions under which the trajectory groups discovered using mixture models for life-course data represent distinct subpopulations as opposed to trajectories of support for an underlying distribution. A logical place to start would be to develop a precise definition of what is meant by distinct groups. We propose the following definition:

Definition 1 (a distinct group): A life-course trajectory group identified by a mixture model is distinct from the rest of the population if and only if any individual who is not already a member of the group cannot become a member of it via small continuous changes (including those in the time schedule) to his or her trajectory.

We suggest that a life-course trajectory group identified by a mixture model is distinct from the rest of the population if and only if it meets the criteria given in Definition 1. In other words, the group should be well separated from the rest of the population in the multivariate space of model parameters. (Note that our use of the word “distinct” is different from and much more precise than the common use of this word in the applied literature on mixture models for life-course data.) This should, of course, be conditional on a good fit of the model to the data in the absolute sense.

For example, identified groups can be declared distinct when there exists a time interval during which the 95% credible bands for predicted individual trajectories show no overlap [e.g., the most likely case for socialization trajectories in open source software projects (Qureshi & Fang 2011)]. Another example of a distinct group is a situation in which the group members follow closely a growth pattern that differs from those of the rest of the population. Thus, for the Montreal opposition data (**Figure 3**), the group of people whose trajectories show a steady decline in opposition during the observational interval would be considered distinct if the variability around the group mean trajectory were small.

Although visual examination of appropriate graphics can usually provide useful insights for the identification of distinct groups, ample room remains for the development of statistical methodology capable of validating these identifications. For example, sensitivity analyses could be used to check the degree to which latent groups are robust to certain changes in study design. If the nature and/or number of groups is not robust to minor changes in study design features (i.e., the

model is not invariant to changes in study design), latent groups may serve only as an approximative device and may not represent distinct subpopulations. Such sensitivity analyses would be consistent with Bartholomew (1995, p. 9), who argued the following in his Presidential Address to the Royal Statistical Society:

The value of science lies in its ability to uncover stable relationships which persist in space and time. If the statistical approach to complex systems is to do more than simply to record the passing scene it also must seek to identify and extract law-like relationships. Invariants, i.e., patterns which persist, are the basis for forecasting, control and insight.

For cases in which there exists strong substantive theory about population groups, mixture modeling of life-course trajectories can be used in a confirmatory manner. The researcher should be encouraged to rely on the theory to specify the number and shapes of the groups *ex ante*, using, for example, constrained splines or latent basis trajectory specifications (Rolfé et al. 2011). If the analysis disconfirms a strong *ex ante* theoretical specification of a number and shapes of groups by revealing more groups, researchers should avoid the temptation to proceed to name and interpret them as if they had been specified *ex ante*. For example, virtually all tests of Moffitt's (1993) *ex ante* dual taxonomy theory have failed to find just two distinct groups. However, instead of concluding that the dual taxonomy theory is unsupported—which would be accurate—researchers invariably assign to the groups *ad hoc* substantive labels and *post hoc* interpretations.

In the absence of unusually strong theory, and if there is no clear indication that model-based latent trajectory groups are distinct (i.e., well-separated in the multivariate parameter space), researchers should treat these groups as trajectories of support approximating the underlying complex distribution of trajectories. We recommend that the researcher refrain from giving descriptive labels to the groups to reduce the likelihood of group reification by readers. We agree with Sterba et al. (2012) that merely adding a caveat sentence in the conclusions after naming the groups is not enough to avoid reification.

To summarize, we have four recommendations. First, we strongly recommend that researchers begin with explicit substantive questions and choose analysis methods that are most suitable for answering them. In some cases, hierarchical linear models may be the most suitable (Raudenbush 2005). However, in other cases, such as those in which individual variability appears to be in intensity and timing and the basic shape of the process can be conceptualized as unimodal, the UCR model can be a natural alternative (Telesca et al. 2012). Similarly, when individual trajectories appear to be situated along a continuum between extreme profiles, researchers might consider estimating extreme profile trajectories and individual-specific mixed memberships scores (Manrique-Vallier 2010, Connor 2006). Researchers might also consider models that allow for a mixture distribution on the underlying continuous factors when supported by existing theory (Slaughter et al. 2009). Second, we recommend using exploratory data analysis for modeling individual trajectories. In particular, we suggest using graphical plots of observed individual trajectories to examine variability among them. Third, researchers should carry out appropriate graphical assessments of the model fit. For example, for continuous outcomes fitted with mixture models, one may plot observed individual trajectories by group. In addition, a researcher might indicate confidence bands for individual predictions around the group means to assess model fit and variability attributable to various sources. Fourth, researchers should carefully examine graphics to assess whether the population contains distinct (i.e., well-separated) groups of life-course trajectories. Only if the groups can be declared as distinct should researchers proceed with assigning group names and carrying out further analyses to prove construct validity and test for group-specific effects (e.g., for an intervention).

In making these recommendations, we aim to help address the main challenge for life-course

research: to approach data analysis by first making conscious decisions about what distinctive features of the data and what types of summaries would be most useful for advancing the field and to then use statistical methodology in a manner that aligns closely with those decisions. Thus, in criminology, where we focus on “breaking bad,” we can ask: When? How bad? For how long?

DISCLOSURE STATEMENT

The authors are not aware of any employment or paid consultancies, board memberships, professional affiliations, or membership in advocacy organizations, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This research was partly supported by the funding from the National Institute on Drug Abuse (R01: DA019148-01A1), a Seed Grant from the Center for Statistics and the Social Sciences (with funds from the University Initiatives Fund at the University of Washington), and the Blumstein-Jordan Endowed Professorship in Sociology. The authors thank Maria Grigoryeva, Blaine Robbins, and Cindy Cai for their assistance in reviewing articles on mixture models for life-course data, and Derek Kreager for his earlier involvement with the Montreal data analysis. We also thank three anonymous reviewers for helpful comments on an earlier draft.

LITERATURE CITED

- Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, ed. BN Petrov, F Csaki, pp. 267–81. Budapest: Akad. Kiado.
- Bandeem-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ. 1997. Latent variable regression for multiple discrete outcomes. *J. Am. Stat. Assoc.* 92(440):1375–86
- Barnett A, Blumstein A, Farrington DP. 1987. Probabilistic models of youthful criminal careers. *Criminology* 25(1):83–108
- Bartholomew DJ. 1995. What is statistics? *J. R. Stat. Soc. A* 158:1–20
- Bauer DJ. 2007. Observations on the use of growth mixture models in psychological research. *Multiv. Behav. Res.* 42(4):757–86
- Bauer DJ, Curran PJ. 2003a. Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychol. Methods* 8(3):338–63
- Bauer DJ, Curran PJ. 2003b. Overextraction of latent trajectory classes: Much ado about nothing? Reply to Rindskopf (2003), Muthén (2003), and Cudeck and Henly (2003). *Psychol. Methods* 8:384–92
- Bauer DJ, Curran PJ. 2004. The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychol. Methods* 9(1):3–29
- Blokland AA, Nagin D, Nieuwebeerta P. 2005. Life span offending trajectories of a Dutch conviction cohort. *Criminology* 43(4):919–54
- Blumstein A, Cohen J. 1979. Estimation of individual crime rates from arrest records. *J. Crim. Law Criminol.* 70(4):561–85
- Blumstein A, Cohen J. 1987. Characterizing criminal careers. *Science* 237(4818):985–91
- Blumstein A, Farrington DP, Moitra S. 1985. Delinquency careers: Innocents, desisters, and persisters. *Crime Justice* 6:187–219
- Blumstein A, Moitra S. 1980. The identification of “career criminals” from “chronic offenders” in a cohort. *Law Policy* 2(3):321–34
- Boscardin CK, Muthén B, Francis DJ, Baker EL. 2008. Early identification of reading difficulties using heterogeneous developmental trajectories. *J. Educ. Psychol.* 100(1):192
- Bowers AJ, Spratt R. 2012. Examining the multiple trajectories associated with dropping out of high school: a growth mixture model analysis. *J. Educ. Res.* 105(3):176–95

- Brame R, Mulvey EP, Piquero AR. 2001. On the development of different kinds of criminal activity. *Sociol. Methods Res.* 29(3):319–41
- Brame R, Nagin DS, Wasserman L. 2006. Exploring some analytical characteristics of finite mixture models. *J. Quant. Criminol.* 22(1):31–59
- Brame R, Paternoster R, Piquero AR. 2012. Thoughts on the analysis of group-based developmental trajectories in criminology. *Justice Q.* 29(4):469–90
- Britt CL. 1992. Constancy and change in the US age distribution of crime: a test of the “invariance hypothesis.” *J. Quant. Criminol.* 8(2):175–87
- Brumback LC, Lindstrom MJ. 2004. Self modeling with flexible, random time transformations. *Biometrics* 60(2):461–70
- Bushway S, Sweeten G, Nieuwbeerta P. 2009. Measuring long term individual trajectories of offending using multiple methods. *J. Quant. Criminol.* 25(3):259–86
- Bushway SD, Thornberry TP, Krohn MD. 2003. Desistance as a developmental process: a comparison of static and dynamic approaches. *J. Quant. Criminol.* 19(2):129–53
- Chassin L, Flora DB, King KM. 2004. Trajectories of alcohol and drug use and dependence from adolescence to adulthood: the effects of familial alcoholism and personality. *J. Abnorm. Psychol.* 113(4):483
- Chib S. 1992. Bayesian inference in the Tobit censored regression model. *J. Econom.* 51:79–99
- Colder CR, Mehta P, Balanda K, Campbell RT, Mayhew K, et al. 2001. Identifying trajectories of adolescent smoking: an application of latent growth mixture modeling. *Health Psychol.* 20(2):127–35
- Connell AM, Frye AA. 2006. Growth mixture modelling in developmental psychology: overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour. *Infant Child Dev.* 15(6):609–21
- Connor JT. 2006. *Multivariate Mixture Models to Describe Longitudinal Patterns of Frailty in American Seniors*. PhD Thesis, Dep. Stat., H. John Heinz III Sch. Public Policy Manag. Carnegie Mellon Univ., Pittsburgh, PA
- De Boor C. 1978. *A Practical Guide to Splines*. Berlin: Springer-Verlag
- D’Unger AV, Land KC, McCall PL. 2002. Sex differences in age patterns of delinquent/criminal careers: results from Poisson latent class analyses of the Philadelphia cohort study. *J. Quant. Criminol.* 18(4):349–75
- Eggleson EP, Laub JH, Sampson RJ. 2004. Methodological sensitivities to latent class analysis of long-term criminal trajectories. *J. Quant. Criminol.* 20(1):1–26
- Elliott MR, Gallo JJ, Ten Have TR, Bogner HR, Katz IR. 2005. Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction. *Biostatistics* 6(1):119–43
- Esbensen F-A, Huizinga D. 1990. Community structure and drug use: from a social disorganization perspective. *Justice Q.* 7:691–709
- Geisser S. 1993. *Predictive Inference: An Introduction*. London: Chapman & Hall
- Gervini D, Gasser T. 2004. Self-modelling warping functions. *J. R. Stat. Soc. B* 66(4):959–71
- Gibbons R, Hedeker D, DuToit S. 2010. Advances in analysis of longitudinal data. *Annu. Rev. Clin. Psychol.* 19:79–107
- Goldstein H. 2011. *Multilevel Statistical Models*. Chichester, UK: Wiley
- Greenberg DF. 1991. Modeling criminal careers. *Criminology* 29(1):17–46
- Greenwood PW, Turner S. 1987. Selective incapacitation revisited. *Why High-Rate Offenders Are Hard to Predict*. Santa Monica, CA: Rand
- Hamil-Luker J, Land KC, Blau J. 2004. Diverse trajectories of cocaine use through early adulthood among rebellious and socially conforming youth. *Soc. Sci. Res.* 33(2):300–21
- Haviland AM, Nagin DS. 2005. Causal inferences with group based trajectory models. *Psychometrika* 70(3):557–78
- Heckman J, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52(2):271–320
- Hirschi T, Gottfredson MR. 1983. Age and the explanation of crime. *Am. J. Sociol.* 89(3):552–48
- Hoeksma JB, Kelderman H. 2006. On growth curves and mixture models. *Infant. Child Dev.* 15(6):627–34

- Hynes K, Clarkberg M. 2005. Women's employment patterns during early parenthood: a group-based trajectory analysis. *J. Marriage Fam.* 67(1):222–39
- Jennings WG, Reingle JM. 2012. On the number and shape of developmental/life-course violence, aggression, and delinquency trajectories: a state-of-the-art review. *J. Crim. Justice* 40(6):472–89
- Jones BL, Nagin DS. 2007. Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociol. Methods Res.* 35(4):542–71
- Jones BL, Nagin DS, Roeder K. 2001. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociol. Methods Res.* 29(3):374–93
- Kass RE, Raftery AE. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90(430):773–95
- Kass RE, Wasserman L. 1995. A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *J. Am. Stat. Assoc.* 90(431):928–34
- Kreuter F, Muthén B. 2008. Analyzing criminal trajectory profiles: bridging multilevel and group-based approaches using growth mixture modeling. *J. Quant. Criminol.* 24(1):1–31
- Lacourse E, Côté S, Nagin DS, Vitaro F, Brendgen M, Tremblay RE. 2002. A longitudinal-experimental approach to testing theories of antisocial behavior development. *Dev. Psychopathol.* 14(4):909–24
- Laird NM, Ware JH. 1982. Random-effects models for longitudinal data. *Biometrics* 38:963–74
- Land KC, Nagin DS, McCall PL. 2001. Discrete-time hazard regression models with hidden heterogeneity: the semiparametric mixed Poisson regression approach. *Sociol. Methods Res.* 29(3):342–73
- Laub JH, Sampson RJ. 2001. Understanding desistance from crime. *Crime Justice* 28:1–69
- Leiby BE, Sammel MD, Ten Have TR, Lynch KG. 2009. Identification of multivariate responders and non-responders by using Bayesian growth curve latent class models. *J. R. Stat. Soc. C* 58(4):505–24
- Leoutsakos J-MS, Muthén BO, Breitner J, Lyketsos CG. 2012. Effects of non-steroidal anti-inflammatory drug treatments on cognitive decline vary by phase of pre-clinical Alzheimer disease: findings from the randomized controlled Alzheimer's disease Anti-inflammatory Prevention Trial. *Int. J. Geriatr. Psychiatry* 27(4):364–74
- Li F, Duncan TE, Hops H. 2001. Examining developmental trajectories in adolescent alcohol use using piecewise growth mixture modeling analysis. *J. Stud. Alcohol Drugs* 62(2):199–210
- Lu ZL, Zhang Z, Lubke G. 2011. Bayesian inference for growth mixture models with latent class dependent missing data. *Multiv. Behav. Res.* 46(4):567–97
- Manrique-Vallier D. 2010. *Longitudinal Mixed Membership Models with Applications to Disability Survey Data*. PhD Thesis, Carnegie Mellon Univ., Pittsburgh, PA
- McLachlan GJ, Peel D. 2000. *Finite Mixture Models*. New York: Wiley
- Moffitt TE. 1993. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol. Rev.* 100:674–701
- Muthén B. 2004. Latent variable analysis. In *The Sage Handbook of Quantitative Methodology for the Social Sciences*, ed. D Kaplan, pp. 345–68. Thousand Oaks, CA: Sage
- Muthén B, Asparouhov T. 2008. Growth mixture modeling: analysis with non-Gaussian random effects. In *Advances in Longitudinal Data Analysis*, ed. G Fitzmaurice, M Davidian, G Verbeke, G Molenberghs, pp. 143–65. Boca Raton, FL: Chapman & Hall/CRC Press
- Muthén B, Brown CH, Masyn K, Jo B, Khoo S-T, et al. 2002. General growth mixture modeling for randomized preventive interventions. *Biostatistics* 3(4):459–75
- Muthén B, Brown HC. 2009. Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Stat. Med.* 28(27):3363–85
- Muthén B, Shedden K. 1999. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 55(2):463–69
- Muthén LK, Muthén B. 2005. *Growth modeling with latent variables using Mplus*. Multilevel Model. in Mplus, Univ. Calif. Los Angel. Stat. Consult. Group, Los Angeles, CA. http://www.ats.ucla.edu/stat/mplus/seminars/mlmMplus_JH/Hopkins_Day1_fixed_May20051.pdf
- Muthén LK, Muthén BO. 2012. *Mplus: Statistical Analysis with Latent Variables: Users's Guide*. Los Angeles, CA: Muthén & Muthén. 7th ed.
- Nagin DS. 2005. *Group-Based Modeling of Development*. Cambridge, MA: Harvard Univ. Press
- Nagin DS, Land KC. 1993. Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poisson model. *Criminology* 31(3):327–62

- Nagin DS, Odgers CL. 2010a. Group-based trajectory modeling in clinical research. *Annu. Rev. Clin. Psychol.* 6:109–38
- Nagin DS, Odgers CL. 2010b. Group-based trajectory modeling (nearly) two decades later. *J. Quant. Criminol.* 26(4):445–53
- Nagin DS, Tremblay RE. 1999. Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Dev.* 70(5):1181–96
- Nagin DS, Tremblay RE. 2001. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol. Methods* 6(1):18–34
- Nagin DS, Tremblay RE. 2005a. Developmental trajectory groups: fact or a useful statistical fiction? *Criminology* 43(4):873–904
- Nagin DS, Tremblay RE. 2005b. What has been learned from group-based trajectory modeling? Examples from physical aggression and other problem behaviors. *Ann. Am. Acad. Polit. Soc. Sci.* 602:82–117
- Neelon B, Swamy GK, Burgette LF, Miranda ML. 2011. A Bayesian growth mixture model to examine maternal hypertension and birth outcomes. *Stat. Med.* 30(22):2721–35
- Odgers CL, Moffitt TE, Broadbent JM, Dickson N, Hancox RJ, et al. 2008. Female and male antisocial trajectories: from childhood origins to adult outcomes. *Dev. Psychopathol.* 20(02):673–716
- Paternoster R, Brame R, Farrington DP. 2001. On the relationship between adolescent and adult conviction frequencies. *J. Quant. Criminol.* 17(3):201–5
- Pettit L. 1990. The conditional predictive ordinate for the normal distribution. *J. R. Stat. Soc. B* 52:175–84
- Pickles A, Croudace T. 2010. Latent mixture models for multivariate and longitudinal outcomes. *Stat. Methods Med. Res.* 19(3):271–89
- Pinheiro JC, Bates DM. 2000. *Mixed-Effects Models in S and S-Plus*. New York: Springer-Verlag
- Piquero AR. 2008. Taking stock of developmental trajectories of criminal activity over the life course. In *The Long View of Crime: A Synthesis of Longitudinal Research*, ed. AM Liberman, pp. 23–78. New York: Springer
- Piquero AR, Blumstein A, Brame R, Haapanen R, Mulvey EP, Nagin DS. 2001. Assessing the impact of exposure time and incapacitation on longitudinal trajectories of criminal offending. *J. Adolesc. Res.* 16(1):54–74
- Piquero AR, Brame R, Mazerolle P, Haapanen R. 2002. Crime in emerging adulthood. *Criminology* 40(1):137–70
- Qureshi I, Fang Y. 2011. Socialization in open source software projects: a growth mixture modeling approach. *Organ. Res. Methods* 14(1):208–38
- Ramsay JO, Li X. 1998. Curve registration. *J. R. Stat. Soc. B* 60:351–63
- Raudenbush SW. 2001. Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annu. Rev. Psychol.* 52:501–25
- Raudenbush SW. 2005. How do we study “What happens next”? *Ann. Am. Acad. Polit. Soc. Sci.* 602:131–44
- Raudenbush SW, Bryk AS. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage
- Reinecke J, Seddig D. 2011. Growth mixture models in longitudinal research. *Adv. Stat. Anal.* 95(4):415–34
- Roeder K, Lynch KG, Nagin DS. 1999. Modeling uncertainty in latent class membership: a case study in criminology. *J. Am. Stat. Assoc.* 94:766–76
- Rolfe MI, Mengersen KL, Vearncombe KJ, Andrew B, Beadle GF. 2011. Bayesian estimation of extent of recovery for aspects of verbal memory in women undergoing adjuvant chemotherapy treatment for breast cancer. *J. R. Stat. Soc. C* 60(5):655–74
- Rousseau J, Mengersen K. 2011. Asymptotic behaviour of the posterior distribution in overfitted mixture models. *J. R. Stat. Soc. B* 73(5):689–710
- Sampson RJ, Laub JH. 1993. *Crime in the Making: Pathways and Turning Points Through the Life Course*. Cambridge, MA: Harvard Univ. Press
- Sampson RJ, Laub JH. 2003. Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology* 41(3):555–92
- Sampson RJ, Laub JH. 2005. A life-course view of the development of crime. *Ann. Am. Acad. Polit. Soc. Sci.* 602(1):12–45
- Saunders JM. 2010. Understanding random effects in group-based trajectory modeling: an application of Moffitt's developmental taxonomy. *J. Drug Issues* 40(1):195–220

- Schwarz G. 1978. Estimating the dimension of a model. *Ann. Stat.* 6(2):461–64
- Skardhamar T. 2010. Distinguishing facts and artifacts in group-based modeling. *Criminology* 48(1):295–320
- Slaughter JC, Herring AH, Thorp JM. 2009. A Bayesian latent variable mixture model for longitudinal fetal growth. *Biometrics* 65(4):1233–42
- Sobel ME, Muthén B. 2012. Compliance mixture modelling with a zero-effect complier class and missing data. *Biometrics* 68(4):1037–45
- Sterba SK, Baldasaro RE, Bauer DJ. 2012. Factors affecting the adequacy and preferability of semiparametric groups-based approximations of continuous growth trajectories. *Multiv. Behav. Res.* 47(4):590–634
- Tarpey T. 2007. Linear transformations and the k -means clustering algorithm: applications to clustering curves. *Am. Stat.* 61(1):34–40
- Tarpey T, Petkova E, Lu Y, Govindarajulu U. 2010. Optimal partitioning for linear mixed effects models: applications to identifying placebo responders. *J. Am. Stat. Assoc.* 105(491):968–77
- Telesca D, Erosheva E, Kreager DA, Matsueda RL. 2012. Modeling criminal careers as departures from a unimodal age-crime curve: the case of marijuana use. *J. Am. Stat. Assoc.* 18(2):159–77
- Telesca D, Inoue LYT. 2008. Bayesian hierarchical curve registration. *J. Am. Stat. Assoc.* 103(481):328–39
- Tremblay RE, Vitaro F, Nagin D, Pagani L, Séguin JR. 2003. The Montréal Longitudinal and Experimental Study: rediscovering the power of descriptions. In *Taking Stock of Delinquency: An Overview of Findings from Contemporary Longitudinal Studies*, ed. TP Thornberry, MD Krohn, pp. 205–54. New York: Kluwer Academic/Plenum
- Underwood MK, Beron KJ, Rosen LH. 2009. Continuity and change in social and physical aggression from middle childhood through early adolescence. *Aggress. Behav.* 35(5):357–75
- Wang C-P, Hendricks Brown C, Bandeen-Roche K. 2005. Residual diagnostics for growth mixture models: Examining the impact of a preventive intervention on multiple trajectories of aggressive behavior. *J. Am. Stat. Assoc.* 100(471):1054–76
- Wiesner M, Capaldi DM. 2003. Relations of childhood and adolescent factors to offending trajectories of young men. *J. Res. Crime Delinq.* 40(3):231–62
- Wiesner M, Windle M. 2004. Assessing covariates of adolescent delinquency trajectories: A latent growth mixture modeling approach. *J. Youth Adolesc.* 33(5):431–42
- Wolfgang M, Figlio R, Sellin T. 1972. *Delinquency in a Birth Cohort*. Chicago: Univ. Chicago Press