Modeling Complex Longitudinal Data from Heterogeneous Samples Using Longitudinal Latent Profile Analysis

Veronica T. Cole

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology.

Chapel Hill
2014

Approved by:

Daniel J. Bauer

Patrick J. Curran

David Thissen

ABSTRACT

Veronica T. Cole: Modeling Complex Longitudinal Data from Heterogeneous Samples
Using Longitudinal Latent Profile Analysis.
(Under the direction of Daniel J. Bauer)

Traditional approaches for examining longitudinal data tend to assume that (1) each individual's shape of change over time follows a pre-specified functional form, and (2) this same basic shape applies to all individuals under study. However, these assumptions are not always tenable in current psychological research. The current report introduces longitudinal latent profile analysis (LLPA), a mixture model which relaxes these assumptions by allowing flexible representation of both inter-individual difference and intra-individual change over time. The LLPA framework allows for this flexibility by modeling the shape of change as a function of two parameters – a time-invariant level, and a vector of time-specific deviations from that overall level – and allowing these paramters to vary categorically between individuals according to latent classes. LLPA and an extension including random effects within-class are applied to an empirical dataset concerning the development of depression in early adulthood, and results are compared to a number of traditional models. The sensitivity of LLPA to random noise in the data is then explored through a brief proof-of-concept simulation. Potential opportunities brought to bear by LLPA, as well as limitations of this approach, are discussed.

ACKNOWLEDGMENTS

First, I would like to thank my family –my mother, my father, Molly, and Marion– for their unyielding support throughout this and all other endeavors. I am particularly indebted to my advisor, Daniel Bauer, who has been a constant source of not only encouragement and support, but also wisdom and vision. His input has made this work, as well as my graduate career thus far, something of which I can truly be proud. I am also grateful to the members of my committee, Patrick Curran and David Thissen, whose feedback and insight have strengthened this thesis immeasurably, and who have always encouraged me to think critically about this work. Bud MacCallum has also offered a number of helpful comments that have guided me in my thinking about this work; I am very thankful for his willingness to share his knowledge with me. Finally, I want to thank my fellow students in the L.L. Thurstone Psychometric Laboratory, who have provided me with support, a rigorous and challenging environment, and most of all friendship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Researchers are often interested in how behaviors or constructs change and develop over time. As such, recent years have seen substantial development of methods to test longitudinally oriented research questions, such as the multilevel model (MLM; Raudenbush and Bryk, 2002; Snijders and Bosker, 2011) and latent curve model (LCM; Bollen and Curran, 2006). These methods allow for random variation in the factors that influence change in different individuals – meaning that each subject is characterized by a (potentially) unique set of parameters governing the initial level of the phenomenon under study (i.e., a random intercept) and change in this phenomenon to time (i.e., a random slope).

While linear growth is specified using a simple intercept and slope, growth factors can take many forms, corresponding to the hypothesized shape of growth over time. For instance, the volume of myelinated neurons in the brain, known as white matter, is known to increase rapidly through childhood and adolescence; this increase continues at a slower rate until middle age, after which white matter volume decreases (Bartzokis et al., 1999; Sowell et al., 2004). Thus, the relationship between age and volume of white matter in the brain resembles an inverted U with its peak around age 50, and is likely better accounted for by a quadratic trajectory than a linear one; in this scenario, a parameter corresponding to quadratic growth can easily be added to the latent curve model. More complex nonlinear forms may be accommodated in a number of ways, from transforming the dependent variable (e.g., Choi, Harring, and Hancock, 2009), to constraining the value of nonlinear parameters to allow them to be modeled as linear terms (e.g., Blozis and Cudeck, 1999), to treating the nonlinear parameters as latent variables (Browne, 1993; Browne and Du Toit, 1991).

Despite the flexibility of the LCM framework, most extensions of these models make two assumptions about the nature of heterogeneity in the data: (1) that inter-individual differences in growth parameters are normally distributed around a population mean; and (2) that intra-individual change follow a prescribed functional form, usually linear or some lower-order polynomial. Thus, these methods are helpful in describing processes in which the same functional form is thought to describe the shape of change for all individuals, per the first assumption, and the basic shape of this functional form is known, per the second assumption.

The main goal of this thesis is to introduce a modeling framework, longitudinal latent profile analysis (LLPA), which allows for the modeling of trajectories without being subject to either of the above assumptions governing the nature of inter-individual differences or the shape of intra-individual change. To contextualize these developments, I first review the rationale for methods that relax the first assumption – specifically, mixture models, which allow categorically different trajectories for different members of the sample. I then introduce LLPA as a minimally parameterized mixture model, and show how it may be used to make inferences about the nature of change at the whole-sample level. Then I extend LLPA's potential utility to individual-level analyses, and thus attempt to show how LLPA may make a novel contribution to person-centered research. Finally, I will compare LLPA to modeling methods which also allow fewer assumptions about the nature of intra- and inter-individual difference in the data.

1.1 Longitudinal mixture models in person-centered analysis

Most extensions of the LCM framework assume that the same general shape of change applies to everyone in the sample. One example of a scenario in which such an assumption might be tenable is the age-white matter curve described earlier: virtually everyone in the population likely experiences the same shape, with white matter volume increasing until middle age and decreasing thereafter. While certain factors, such as neuropsychiatric illnesses, may alter the rate of change in white matter volume (Bartzokis et al., 2003) the fundamental shape of the growth curve across the lifespan is probably quadratic for each individual.

In psychology, however, researchers are frequently interested in phenomena that may follow trajectories of completely different shapes for different individuals in the sample. A broad class of models that can be implemented for evaluating this heterogeneity is mixture models (McLachlan and Peel, 2000). Though a formal treatment of mixture models will be given later, the basic premise of these models is that they treat subjects as though they come from distinct subpopulations (known as classes or components), allowing for statistical inference to be made on the basis of these subpopulations. Originally conceived as a method of classifying natural phenomena (Pearson, 1894), mixtures have gained use in the social sciences, with both cross-sectional (e.g., Lazarsfeld and Henry, 1968) and longitudinal (e.g., Muthén and Shedden, 1999; Nagin, 1999) applications.

Titterington, Smith, and Makov (1985) distinguish between direct and indirect applications of mixture modeling. The vast majority of published applications of longitudinal mixture models are direct applications, in which the goal is to ascertain subgroup membership – in these models, categorical variation is a substantive outcome of interest. Direct apppications are consistent with the search for meaningful typologies of psychological phenomena, a goal which has characterized much psychological work throughout the past century (e.g., Meehl, 1992). In indirect applications, by contrast, parameters from the mixture model are usually used to ascertain information about the population as a whole, aggregating across clusters; thus, in indirect applications, the categorical latent variable is typically not an outcome of intrinsic interest.

Both direct and indirect applications of mixture models have a great deal of potential relevance to a ”person-centered” framework (Bergman and Magnusson, 1997). Rather than isolating the relationship between one variable and another, a person-centered approach seeks to isolate individuals, or clusters thereof, who show particular patterns of response to any number of variables, including both time-varying indicators (e.g., substance abuse at time t) and covariates. The draw of such an approach is that it allows individuals to be characterized holistically in terms of interactions between any number of internal and external factors working in concert (Bauer and Shanahan, 2007). Proponents argue that this approach is far closer to the reality of complicated developmental processes than variable-centered techniques, which are often applications of regression models emphasizing single factors' effects controlling for all others.

One substantive area in which person-centered analyses have frequently taken the form of direct applications of mixture models is the developmental etiology of substance abuse. For instance, Schulenberg et al. (1996a) find six trajectories of binge drinking between ages 18 and 24, and link these trajectories post-hoc to a number of covariates (1996b). Their work shows a number of individuals maintaining habits of either minimal, moderate, or heavy binge drinking, some decreasing linearly in the frequency of binge drinking, and others (termed ”fling” binge drinkers by the authors) showing a period of increased binge drinking toward the middle of the study period, after which drinking decreases substantially, thus resembling a quadratic trend (Maggs and Schulenberg, 2005). Casswell, Pledger and Pratrap (1996) report a similar quadratic pattern in one subset of their sample, only these individuals experienced peak binge drinking around age 26. Similar findings of heterogeneity have also been reported when considering younger samples (e.g.,King et al., 2000; Chassin, Pitts, and Prost, 2002; Sher, Jackson, and Steinley, 2011), as well as substances besides alcohol such as marijuana (e.g, Schulenberg et al., 2005; White, Labouvie, and Papadaratsakis, 2005; Windle and Wiesner, 2004), and cigarettes (e.g., White, Padina, and Chen, 2002).

These analyses, which focus on decomposing a heterogeneous sample into ”prototypical” trajectories of substance abuse and relating these trajectories to covariates, still assume that each group is defined by a parametric form. It is possible, however, that assuming linear or quadratic increases in substance use might be an oversimplification of what is occurring in the data. For instance, the ”fling” group of binge drinkers identified by Schulenberg et al. (1996a; 1996b) may not uniformly show a single period of increased binge drinking around college age, as would be implied by the quadratic trajectory found by the researchers. Perhaps some subjects show this trend of one isolated period of heavy drinking during college but, additionally, there are individuals in the sample who show multiple recurring periods of increased drinking, as well as some who show a single period of binge drinking past college age, et cetera. Furthermore, it may be the case that these sub-classifications of ”fling” drinkers are differentiated from one another by covariates (e.g., Wechsler et al., 1995; Weitzman, Nelson, Wechsler, 2003). Therefore, imposing functional forms on trajectories of binge drinking across time might sacrifice valuable information required for a truly person-centered approach, both by oversimplifying the true shape of change as it occurs on an individual level, and by obscuring the relationship between these changes and contextual factors.

More generally, the direct application of mixture models is a controversial practice for a number of reasons. On a conceptual level, direct applications yield subgroups, which are often reified in the literature as ”types” that exist as real entities. This can be particularly problematic because mixture models are likely to spuriously extract multiple classes in the sample even if the data come from a single population. This can happen when the data show even mild departures from normality (Bauer and Curran, 2003a; 2003b), or, in the multivariate case, if the indicators are highly correlated; thus, in most longitudinal applications the extraction of multiple classes is more likely than not, regardless of the underlying nature of inter-individual differences (Bauer and Curran, 2004; Bauer, 2007). However, some of the features of mixture models that make their direct application problematic may actually make them well-suited to indirect application (Bauer, 2005). Pursuing this possibility, I will now introduce a new mixture model, longitudinal latent profile analysis (LLPA), and discuss several potential indirect applications of this model.

1.2 Longitudinal Latent Profile Analysis

Latent profile analysis (LPA; Gibson, 1959) is a type of mixture modeling that was developed for and has traditionally been applied to data obtained at a single time point on multiple, distinct continuous indicators. Direct applications of LPA are common, and LPA has shown its utility in studies seeking to classify patients into subgroups of symptom profiles, with degree of endorsement of each symptom being an indicator (e.g., Mitchell et al., 2007; Holliday et al., 2009). LPA has also been used in community samples with the aim of determining the level and nature of certain psychiatric issues in the population (e.g., Wade, Crosby and Martin, 2006).

In a basic LPA, the observed score for individual *i* (*i*=1,...,*N*) on indicator *j* given membership in class *k* is given by:

$$y_{ij}^{(k)} = \xi_j^{(k)} + E_{ij} \tag{1}$$

where $\xi_j^{(k)}$ is the true score on indicator *j* for class *k*. Here error is represented by $E_{ij} \sim N(0, \sigma_j^2)$, where σ_j^2 is the residual, or within-class, variance of indicator *j* across individuals. Importantly, residuals are assumed uncorrelated with one another across variables within class. Thus class membership, and therefore between-class differences in true scores, is the only source of covariance between the variables; this is termed the *conditional independence* assumption.

The traditional LPA framework, which forms profiles based on multiple variables at one time, may be adapted to a longitudinal setting, such that profiles capture differences in a single variable over time. Equation 1 may be reconsidered so that *j* indexes occasions of measurement of a single variable, and thus ξ_j represents scores on this variable across time (e.g., Degnan et al., 2010). When applied to longitudinal data, it may be helpful to reparameterize the LPA by decomposing each time-specific indicator into two parameters:

a class-specific grand mean of the indicators, $\mu^{(k)}$, which represents the overall level of the indicator in class *k* across time; and time-specific deviations from the overall level, $\delta_t^{(k)}$, which sum to zero and represent the shape of variation around the within-person mean. The longitudinal latent profile analysis (LLPA) may thus be written as:

$$y_{it}^{(k)} = \mu^{(k)} + \delta_t^{(k)} + E_{it} \tag{2}$$

Here y_{it} refers not to individual *i*'s score on item *j*, as in Equation 1, but to individual *i*'s score at time *t* (t=1,...,*T*), given membership in class *k*. The individual error component is given by $E_{it} \sim N(0, \sigma_t^2)$, where σ_t^2 is the variance at time *t* across individuals. Importantly, there are no constraints on the values of $\delta_t^{(k)}$ aside from the fact that all values for each class must sum to zero – they simply represent vertical deviations from the overall level at time *t*. Therefore, the model allows for maximum flexibility, as no parametric relationship between time and the indicator is modeled.

While $y_{it}^{(k)}$ represents the observed score for individual *i* given membership to class *k*, one may aggregate over the classes using π_k, the proportion of the sample in class *k*, to obtain the marginal means and covariance structure for the whole sample. The model-implied marginal means are:

$$\boldsymbol{\mu}_y = \sum_{k=1}^{K} \pi^{(k)}(\boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}^{(k)}) \tag{3}$$

where $\boldsymbol{\mu}_y$ represents a vector of length *T* containing the estimated marginal means of the repeated measures at each time point; $\boldsymbol{\mu}^{(k)}$ is a vector of length *T* containing the level of class *k*, repeated *t* times; and $\boldsymbol{\delta}_t^{(k)}$ is a vector of length *T* containing time-varying shape parameters for class *k*.

Similarly, the model-implied marginal covariance matrix can be ascertained. Because Equation 2 represents a decomposition of the repeated measure $\boldsymbol{\xi}_j^{(k)}$ into two parts, we will refer to the vector of class-specific means of indicators formed by adding $\boldsymbol{\mu}^{(k)}$ and $\boldsymbol{\delta}^{(k)}$ as $\boldsymbol{\xi}^{(k)}$ for the purpose of brevity; because $\boldsymbol{\xi}^{(k)} = \boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}^{(k)}$, $\boldsymbol{\xi}^{(k)}$ represents the vector of means at each of time *t* for class *k*. Thus covariance matrix of indicators in the sample is calculated as:

$$\Sigma_{yy} = \sum_{k=1}^{K} \pi^{(k)} (\boldsymbol{\xi}^{(k)} - \boldsymbol{\mu}_y)(\boldsymbol{\xi}^{(k)} - \boldsymbol{\mu}_y)' + \sum_{k=1}^{K} \pi^{(k)} \boldsymbol{\Sigma}^{(k)} \tag{4}$$

The first term refers to the between-class contribution to the marginal covariance matrix; the second term refers to within-class contribution to the marginal covariance. The first term models the variances of indicators, as well as covariances between them, as a result of differences in within-class means; in the LLPA, each class' vector of means represents the unique longitudinal profile of that class. Here $\boldsymbol{\Sigma}^{(k)}$ is a diagonal matrix, with the within-class variance of repeated measures in class *k* on the diagonal, and thus the second term of this equation contributes only to the variances of repeated measures, not covariances, in the marginal covariance matrix. Thus, all covariance in the observed scores y_{it} is thought to be explained by the first term, or differences in between-class means. Thus, *conditional independence* assumption, which holds that covariances between indicators are accounted for by class membership, holds as in Equation 1.

Beyond simply describing the patterns of change in the data, researchers often want to use covariates to predict which members of the sample will follow specific patterns. Person-level covariates can readily be incorporated into the LPA, through a multinomial regression sub-model for the class probabilities. Importantly, these predictors affect the model only through their effects on class membership – they do not directly affect the level or shape parameters within-class. In conventional LPA, the addition of both time-invariant and time-varying covariates can be incorporated, and help to identify classes. For instance, Vaughn et al. (2007) find four categories of psychopathology and substance use in juvenile offenders by using subscales of the Brief Symptom Inventory (BSI) as indicators, and information about offense history (e.g., violent vs. nonviolent offending, age at first offense) as covariates. However, the person-level predictors considered in this thesis are thought to be antecedent to the groupings themselves and are thus fundamentally different in meaning from the repeatedly measured class indicator (Lubke and Muthén, 2007; Marsh, Ludtke, Trautwein, and Morin, 2009). For this reason, only time-invariant covariates are considered formally here. Considering only time-invariant predictors in the current analysis helps to make the distinction between covariates and indicators somewhat less hazy: indicators are a time-varying dimension, while covariates are time-invariant antecedents influencing class membership but not the shape of the trajectory in each class.

The class probabilities, dependent on covariates, are defined through a multinomial logistic regression:

$$\pi^{(k)}(\boldsymbol{z}_i) = \frac{\exp(\theta_0^{(k)} + \boldsymbol{\theta}^{(k)\prime}\boldsymbol{z}_i)}{\sum_{k=1}^{K}\exp(\theta_0^{(k)} + \boldsymbol{\theta}^{(k)\prime}\boldsymbol{z}_i)} \tag{5}$$

where $\theta_0^{(k)}$ is a class-specific multinomial intercept for class k, and $\boldsymbol{\theta}^{(k)}$ is the multinomial regression coefficient relating the covariate vector $\boldsymbol{z}_i$ to membership in class k. As in most applications of multinomial logistic regression, one class is treated as a reference class, in which $\theta_0^{(k)}$ and $\theta^{(k)}$ are constrained to zero for model identification.

The vector of model-implied means for the repeated measures given $\boldsymbol{z}_i$ can thus be given by a slightly altered version of Equation 3, which conditions on the covariate vector $\boldsymbol{z}_i$:

$$\boldsymbol{\mu}_{yi} = \sum_{k=1}^{K}\pi^{(k)}(\boldsymbol{z}_i)(\boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}^{(k)}) \tag{6}$$

Given Equations 5 and 6, expected trajectories can be calculated and plotted at different values of the covariates, as is often done in multilevel modeling and latent curve analysis (Preacher, Curran, and Bauer, 2006). This strategy can serve a number of purposes. First, one can generate values of $\pi^{(k)}(\boldsymbol{z}_i)$ for selected values of a single covariate (e.g., low, medium, and high) holding other covariates constant (e.g., at their means) using Equation 5, and then use the resultant value to generate a predicted trajectory using Equation 6. Such results may be useful for determining the nature of the relationship between predictors and trajectories on average in the population. However, LLPA can be extended further to provide insight into the shape of change at the individual level; I now explore the two main ways, using Equations 5 and 6 and extensions thereof, of using LLPA to make individual predictions.

1.3 Individual-level inference in LLPA

Pursuant to the goals of person-centered analysis, it is often of interest to plot model-predicted trajectories for individuals in the sample. Assessment of model fit at the individual level is common in the MLM literature (Rabe-Hesketh and Skrondal, 2008), but relatively rare in mixture models. However, these individual model-implied trajectories can provide information about how well the model approximates the observed trajectories at the individual level.

LLPA can be used to accomplish this goal in two ways. First, one can generate predicted values of $\pi^{(k)}(\boldsymbol{z}_i)$ for each individual in the sample, given each individual's vector of predictors $\boldsymbol{z}_i$ – in other words, rather than using Equations 5 and 6 to calculate hypothetical trajectories given values of covariates chosen by the researcher, one could input the combination of covariates specific to individual *i* to obtain that individual's predicted trajectory.

In this formulation, $\pi^{(k)}(\boldsymbol{z}_i)$ is considered to be the predicted probability of class membership, and is thus used to weight each class' trajectory. However, $\pi^{(k)}(\boldsymbol{z}_i)$ is essentially a *prior* probability in the sense that it doesn't incorporate any information about the vector of observed scores into the probability of individual *i* belonging to class *k*. In order to incorporate information about both covariates $\boldsymbol{z}_i$ and the observed time-varying indicator $\boldsymbol{y}_i$, one may obtain the *posterior* probability of individual *i* belonging to class *k*. Given a vector of length T consisting of all of subject i's observed scores, referred to as $\boldsymbol{y}_i$, this posterior probability is given by:

$$\tau_i^{(k)}(\boldsymbol{z}_i) = \frac{\pi^{(k)}(\boldsymbol{z}_i) f^{(k)}(\boldsymbol{y}_i|\boldsymbol{z}_i)}{f(\boldsymbol{y}_i|\boldsymbol{z}_i)} \tag{7}$$

where $f^{(k)}(\boldsymbol{y}_i|\boldsymbol{z}_i)$ is the class-specific multivariate normal density for $\boldsymbol{y}_i$ conditional on $\boldsymbol{z}_i$; and $f(\boldsymbol{y}_i|\boldsymbol{z}_i)$ is the marginal multivariate normal density for $\boldsymbol{y}_i$ conditional on $\boldsymbol{z}_i$.

Thus, the posterior probability of person i belonging to class k is calculated by weighting the class-specific likelihood by the whole-sample membership proportion for class k given covariates $\boldsymbol{z}$, and dividing by the whole sample likelihood. Importantly, each member of the sample has a nonzero probability of belonging to each class, and all posterior probabilities over the K classes sum to 1 for individual i.

Just as the whole-sample trajectories can be calculated by weighting class-specific trajectories according to class membership proportion (found in Equation 6), individual trajectories can be calculated weighting by an individual's posterior probability of belonging to a particular class. Thus, an individual's model-predicted trajectory is given by:

$$\hat{\boldsymbol{y}}_i = \sum_{k=1}^{K} \tau_i^{(k)}(\boldsymbol{z}_i)(\boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}^{(k)}) \tag{8}$$

Equation 8 indicates that class-specific values of level and shape can be weighted by the posterior probability of membership to each class (given by Equation 7), and aggregated to yield the predicted trajectory for person i. For each individual, this total trajectory represents the weighted sum of all classes' trajectories. Contrasting this trajectory to with one found using Equation 6 represents a distinction similar to that between marginal and conditional predicted trajectories in the mixed models literature (Laird and Ware, 1982). As Equation 6 predicts values for individual i only on the basis of prior probability of group membership, the resultant trajectory is analogous to a marginal predicted value based only on fixed effects. By contrast, Equation 8 represents a conditional predicted trajectory, as it incorporates the equivalent of a random effect in $\tau_i^{(k)}(\boldsymbol{z}_i)$ in weighting the class' trajectories; thus trajectories predicted using Equation 8 with sample values in place of population parameters are empirical Bayes estimates.

Both the marginal and conditional predicted trajectories for individuals in the sample may be of interest, and they may serve distinct but complementary purposes. In particular, gauging the discrepancy between an individual's observed and marginal predicted trajectories may yield an impression of whether a given individual is ”typical” given a certain set of covariates or predisposing factors (e.g., Skrondal and Rabe-Hesketh, 2003). By contrast, the discrepancy between an individual's observed and conditional predicted trajectories will be used to index individual-level model fit (Coffman and Milsap, 2006; Skrondal and Rabe-Hesketh, 2009).

1.4 Random-effects LLPA

It is often of greater interest in longitudinal research to ascertain differences in the shape of a phenomenon, rather than differences in level. However, in the LLPA as it is currently formulated, differences in overall level may have more influence on class membership than shape, even if heterogeneity in shape is the primary outcome of interest. For example, given the vast heterogeneity in overall level of depression among adolescents (e.g., Dekker et al., 2007), one might imagine that classes based on longitudinal measures of depression would largely reflect variance in level, potentially obscuring interesting heterogeneity in shape.

One potential modification to the LLPA that may address this concern is to allow for within-class variability in the level parameter. In this scenario Equation 2 would be altered as follows:

$$y_{it}^{(k)} = [\mu^{(k)} + \alpha_i^{(k)}] + \delta_t^{(k)} + E_{it} \tag{9}$$

where $\alpha_i^{(k)} \sim N(0, \sigma_\alpha^2)$. Whereas the term $\mu^{(k)}$ in Equation 2 represented a constant within-class level parameter, here it is simply the mean value of level for class k. Thus, the portion of Equation 9 in brackets refers to a person-specific level parameter, comprised of mean $\mu^{(k)}$ and person-specific deviation $\alpha_i^{(k)}$, where these deviations are normally distributed across persons within-class.

In this model, referred to here as the random-effects LLPA, fewer classes would likely be needed to approximate heterogeneity in profiles of growth, and the latent class variable would capture more information about the shape of change for the variable, which may produce more substantively interesting and interpretable profiles. Importantly, this also means that covariates, in that they are linked to the variables exclusively through class membership, may now only explain one of two pools of variability in the level parameter. The random-effects LLPA will be further explored in an empirical analysis later in this thesis.

Both with and without within-class random effects for level, LLPA's lack of assumptions about the functional form of growth may help to preserve a greater level of detail about within-person variability than standard extensions of the LCM. However, LLPA is not the only model that has been proposed to facilitate person-centered analysis through flexible modeling of change. In order to better understand the unique features of LLPA, I will now consider its potential advantages and disadvantages relative to two existing techniques allowing different types of within- and between-person variability: longitudinal mixture models, and single-class extensions of the latent growth model allowing different forms of growth.

1.5 Relation to other models

One sort of longitudinal mixture model that has gained popularity in recent years is the semiparametric growth model (SPGM; Nagin and Land, 1993; Nagin, 1999). In an SPGM, subjects are assumed to come from *K* classes, with each class *k* defined by its own growth equation. For instance, a linear SPGM would be of the form:

$$y_{it}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} time_{it} + E_{it} \tag{10}$$

where $\beta_0^{(k)}$ and $\beta_1^{(k)}$ vary between (not within) classes, and $E_{ij} \sim N(0, \sigma_t^2)$.

If time is centered at the midpoint, the intercept term $\beta_0^{(k)}$ Equation 10 represents the within-class average level of the observations, just like the level parameter $\mu^{(k)}$ in Equation 2. Where SPGM differs from LLPA is in the specification of a parametric relationship between the repeated measure and time: whereas the shape parameter $\delta_t^{(k)}$ in Equation 2 does not directly incorporate any parametric function of time, the term $\beta_1^{(k)} time_{it}$ in Equation 10 imposes straight-line growth. While Equation 10 represents linear growth in each class, quadratic or cubic growth can also be modeled (e.g., Karp, O'Loughlin, Paradis, Hanley, and Difranza, 2005; Tucker, Orlando, and Ellickson, 2003), as can more complex functions such as Gompertz and logistic curves (Grimm and Ram, 2009; Grimm, Ram, and Estabrook, 2010). The disadvantage of this approach, relative to LLPA, is that the researcher must know the function a priori, whereas this function may often be unknown or uncertain. An advantage of SPGM over LLPA, however, is that the specification of a parametric form allows for individually-varying times of observation.

Just as the LLPA can be extended to incorporate within-class variance of level, so too can the SPGM be extended to allow within-class variance of growth parameters. Longitudinal mixture models that allow growth parameters to vary within-class are generally referred to as growth mixture models (GMM; Muthén and Shedden, 1999). A GMM of linear growth would be given by:

$$y_{it}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} time_{it} + E_{it} \tag{11}$$

where $\begin{bmatrix} \beta_0^{(k)} \\ \beta_1^{(k)} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{\beta_0}^{(k)} \\ \mu_{\beta_1}^{(k)} \end{bmatrix}, \begin{bmatrix} \sigma_{\beta_0}^{2(k)} & \\ \sigma_{\beta_0\beta_1}^{(k)} & \sigma_{\beta_1}^{2(k)} \end{bmatrix} \right)$.

While growth itself is represented just as in the SPGM, here the growth parameters are assumed to vary normally within class around the class mean; thus $\beta_0^{(k)}$ and $\beta_1^{(k)}$ are essentially random effects within class. Furthermore, one can formulate longitudinal mixture models in which different components have different functional forms of growth, such as one class showing linear growth and another class with quadratic growth (Muthén and Muthén, 2000; Muthén, 2001). Freeing the variance of growth parameters changes their interpretation, a point that has been debated widely within the context of direct applications of mixture models (Nagin, 2005; Nagin and Tremblay, 2005; Muthén, 2006). Rather than being a parameter defining the entirety of the group, in a GMM the parameter value simply represents a mean around which there is random variance, just as in a single-class latent growth model. Importantly, however, these interpretational changes are most relevant in direct applications, given that within-class parameters are typically not interpreted in indirect applications. In this sense, what both GMM and the random-effects LLPA represent is a combination of categorical and continuous treatments of variance in growth parameters, since these growth parameters are allowed to vary categorically according to class membership but also continuously around a class mean parameter value.

The LLPA and other longitudinal mixture models (SPGM, GMM) are similar in that they treat trajectories as varying categorically between subjects and thus relax the standard LCM's assumptions about the nature of *inter*-individual differences. Unlike the SPGM and GMM, however, the LLPA is also minimally parameterized with respect to the shape of change, and thus relaxes assumptions about the nature of intra-individual differences. Another model that shares LLPA's lack of a pre-specified functional form is a LCM with completely freed loadings, sometimes referred to as the latent basis LCM (Meredith and Tisak, 1984, 1990; McArdle and Epstein, 1987). This model is specified as:

$$Y_{it} = \beta_{0i} + \beta_{1i}\lambda_t + E_{it} \tag{12}$$

where λ_t designates a factor loading that (aside from identification restrictions) is estimated from the data. Just as an LLPA is essentially an SPGM without pre-defined functions of time, a latent basis model is an LCM without pre-specified loadings for time, and therefore no pre-defined functional form. Two loadings are fixed at 0 and 1 to set the metric of the factor; the remaining loadings are estimated freely. This is contrasted with models such as a linear LCM, in which the loadings of λ_t are fixed at linearly increasing values, usually [0, 1, 2,..., T-1]; unlike these models, time is unstructured in the latent basis model. Because the loadings of time are freely estimated, the metric of time is not always easily interpreted; McArdle (2004) recommends setting the first and last loadings at 0 and 1 respectively, in which case the loading of each time point represents the proportion of total change that has occurred at time t. This particular parameterization is most useful if change is monotonic.

Though both the loadings and the growth factors lose some degree of interpretability under this approach, the latent basis allows not only for a nonlinear functional form but for the lack of any a priori specification of functional form at all. As with LLPA, this degree of flexibility can be greatly beneficial in modeling change when the functional form of intra-individual change is unknown. For instance, Grimm (2007), modeling the relationship between depression and academic achievement between ages 7 and 14, finds that a latent basis model provides an excellent fit to the data; similarly, Grimm and Ram (2009) find that a latent basis model very well approximates the fluctuations in cortisol over trials of a stress test. However, a disadvantage of the latent basis model is that it assumes that the same basic shape of change is the same for all individuals, with only differences in magnitude or direction.

1.6 Summary and Research Aims

Table 1 summarizes the models considered thus far in terms of their treatment of intra- and inter-individual variability. The current report considers the ways in which LLPA compares with a number of the other models in Table 1 at approximating functional form heterogeneity in longitudinal data. These comparisons are not driven by particular hypotheses, as the goals of these analyses are mostly exploratory. An empirical examination of longitudinal data investigates the various methods' ability to fit trajectories at both the aggregate and individual levels in the 1997 National Longitudinal Survey of Youth (NLSY97). Following this analysis, a brief proof-of-concept simulation is conducted with the goal of elucidating model selection between LLPA and other candidate models.

2 METHODS

2.1 The National Longitudinal Survey of Youth

Data come from the 1997 National Longitudinal Survey of Youth (NLSY97), an ongoing longitudinal study that has collected data on education, employment, government program participation, crime, family life, substance abuse, and health in a nationally representative sample of adolescents. From 1997 to 2000, data were collected at one-year intervals; thereafter, data were collected at two-year intervals. Information sources included face-to-face and computerized interviews with subjects and their parents, as well as examination of school and administrative records. For the current study, I examine data taken from six of the time points: 2000, 2002, 2004, 2006, 2008, and 2010.

From the larger NLSY dataset a smaller sample (N = 1686), consisting exclusively of individuals who were 14 years old at the time of their first interview in 1997, is analyzed. The age requirement ensured that all subjects were transitioning from adolescence to adulthood between 2000 and 2010. The sample was 50.4 percent male, and relatively ethnically diverse, with 26.8, 20.0, .08, and 50.5 percent of respondents identifying as Black, Hispanic/Latino, mixed race, and neither Black nor Hispanic/Latino, respectively. For ease of model fitting, this sample includes individuals who were interviewed at least half of the time points. Patterns of missing data are shown in Table 2.

2.2 Measures

The main outcome of interest, depression, was measured in the NLSY using five items from the Center for Epidemiological Studies Depression scale (CES-D), listed in Table 3. During this part of the interview subjects were asked how often during the past month they had experienced various emotions, and gave a response from 1-4, with a score of 1 representing all of the time and 4 representing none of the time. The data were coded such that higher values were indicative of less depression. An aggregate score was created from these five items, which showed moderately good internal consistency, with Cronbach's alpha of .76, .76, .77, .77, .80, and .81 at each of the successive six time points, respectively.

Four covariates were also included in the conditional analyses. First, given the frequent finding that women tend to report higher levels of depression than men, gender was included as a covariate in the model, coded as 1 if the subject was female and 0 otherwise. Race was also included in the model, coded as 1 if the subject was Caucasian and 0 otherwise. In order to consider theories relating physical and mental health in development (e.g., Repetti, Taylor, and Seeman, 2002) we considered parent ratings of the subject's general health at age 13, two years before the subject's first interview. Parent-rated health was an ordinal variable, coded from 1-5, with lower scores representing better overall health. Finally, given that there are some reports of different clinical patterns among college students and their non-college-attending counterparts (e.g., Gfoerer, Greenblatt, and Wright, 1997; Blanco et al., 2008), class membership was also regressed on a binary variable representing college attendance, coded as 1 if the subject attended college by age 23 and 0 otherwise.

2.3 Analysis

2.3.1 Model Fitting

Models were fit to the data using Mplus 7.0 (Muthén and Muthén, 2012). Table 1 shows all sets of models that were tested, comprising six sets in all. The first three require a priori specification of a functional form: (1) a standard latent curve model (LCM) with linear and quadratic components; (2) a semiparametric growth model (SPGM) with linear and quadratic components; and a (3) growth mixture model (GMM) with fixed slope and quadratic components and an intercept component estimated as a random effect with a normal distribution within class. The final three models do not require a priori specification of functional form: (4) a latent basis model (LBM) with all loadings except the first and last ones freely estimated; (5) a longitudinal latent profile analysis (LLPA) with fixed level and shape factors; and (6) a longitudinal latent profile analysis with a random effect for level within class (LLPA-RE).

2.3.2 Model Comparison at the Global Level

For the mixture models, the Bayesian Information Criterion (BIC; Schwartz, 1978) was used to select the optimal number of classes. The BIC is equal to $-2\times\log(LL)+q(\log(N))$, where *LL* represents the log-likelihood of the model, *q* represents the number of repeated measures, and *N* represents the sample size; a lower BIC thus indicates a stronger fit and more model parsimony at the global level. Both across models and within solutions for the same model (e.g., a 4-class LLPA vs. a 5-class LLPA), BIC was used to compare fit.

2.3.3 Individual Level Analysis

Assessment of fit at the individual level proceeded via the Root Mean Square Residual (RMSR), as defined by Coffman and Millsap (2006). The RMSR directly measures the proximity between individual *i*'s observed trajectory $\boldsymbol{y}_i$, and individual *i*'s model-predicted trajectory. The model-predicted trajectory is defined two ways: first as the marginal predicted trajectory given by Equation 6, $\boldsymbol{\mu}_{yi}$, and second as the conditional predicted trajectory given by Equation 8, $\hat{\boldsymbol{y}}_i$. Thus, the individual fit to the marginal and conditional predicted trajectories will be assessed, and termed $RMSR_{Mi}$ and $RMSR_{Ci}$ respectively and defined as:

$$RMSR_{Mi} = \sqrt{\frac{\sum_{t=1}^{T_i} (y_{it} - \mu_{yit})^2}{T_i}} \tag{13}$$

$$RMSR_{Ci} = \sqrt{\frac{\sum_{t=1}^{T_i} (y_{it} - \hat{y}_{it})^2}{T_i}} \tag{14}$$

where T_i is equal to the number of times at which each individual *i* is assessed.

The RMSR provides an intuitive basis for comparing the fit of predicted trajectories generated using the LLPA, SPGM, and latent basis model to the observed trajectory for each individual. One important caveat to the RMSR is that, unlike the BIC, the RMSR does not take into account any measure of model parsimony; thus, it is exclusively a measure of closeness of fit and not necessarily generalizability.

For the conditional analyses, the RMSR was calculated two ways – first using Equation 14, yielding $RMSR_{Ci}$, the RMSR comparing subject *i*'s observed and posterior-predicted trajectories, and then using Equation 13, yielding $RMSR_{Mi}$, the RMSR comparing subject *i*'s observed and prior predicted trajectories given values of covariates. The RMSR was examined only using equation 14 in the unconditional analysis. The reason for this seeming discrepancy is that, in an unconditional analysis, there are no covariates on which to condition the prior probability of class membership and thus the prior-predicted trajectories are not meaningful at the individual level. First, RMSR values were compared among all of the models, to see if any models provide closer fit to individual data than others.

Then, in order to determine the possibility that models fit differentialy well for individuals with different values of covariates, I examined Pearson's correlation coefficients relating RMSR values ($RMSR_{Ci}$ in both the conditional and unconditional analyses, and $RMSR_{Mi}$ in the conditional analysis) to the four covariates included in the conditional analysis – gender, race, college attendance, and parent-rated health at age 13. Additionally, I examined the correlation between the RMSR under each model and the number of missing CES-D scores to see whether each model showed differentially close fit for individuals with more or less missingness. Finally, in order to demonstrate the graphical capabilities of LLPA to gain exploratory, non-hypothesis-driven insight into individual trajectories, the individual predicted and observed curves for a selection of cases were plotted. The methods of Pek, Bauer and Losardo (2011) were used to generate bootstrap-predicted uncertainty around each individual trajectory. This method entails using the covariance matrix of model parameters to simulate 100 alternative sets of parameters, which are used to generate 100 alternative trajectories for each individual; by plotting these, one can see a random sample of the trajectories that would be expected given the distribution of sampling error in the parameter estimates.

3 RESULTS

3.1 Overall model fit

Table 4shows the BIC values for all unconditional models that were tested, ranked according to overall standing (with lower BIC values indicating better performance and thus being ranked higher). The latent curve and latent basis models performed poorly relative to other models; among these options, however, BIC was lower in the versions of both models in which the residual variances of all time points were assumed equal. The mixture models with only categorical variation between latent classes in growth parameters showed wide variation in the degree of fit. Solutions with more classes fit better, with an 11-class solution fitting best in the SPGM – however solutions with more than 11 classes did not converge, and the best likelihood was replicated in neither the 10- nor the 11-class solution. Thus, a 9-class SPGM is considered in further analyses, as this was the last solution at a probable global maximum. Fit was better in the LLPA than in the SPGM; a 12-class solution was the best fit. The GMM and LLPA-RE, which allow both categorical and continuous variation in growth parameters, fit the data the best overall. In the GMM a 5-class solution fit the data the best. In the LLPA-RE an 11-class solution represented the best balance of fit and parsimony. Importantly, this was the best fitting model in the whole sample, with a BIC value almost 200 points below the best-fitting GMM.

Taken together, these results suggest two trends in the data. First, models allowing both continuous and categorical variation across individuals in growth parameters (the GMM and the LLPA-RE) fit better than models allowing only categorical variation in growth parameters (the SPGM and the LLPA), which in turn (given a large number of classes) fit better than those allowing only continuous inter-individual variation in growth parameters (the latent curve and latent basis models). Second, regardless of the specification of between-person variation in growth parameters, models which do not require pre-specification of functional form (the latent basis model, LLPA, and LLPA-RE) fit better than models that do (latent curve model, SPGM, and GMM).

Findings related to the conditional analysis, which are shown in Table 5 showed a few differences from the unconditional analysis. First, models with fewer classes were generally favored by the BIC in the mixture model solutions, likely reflecting the fact that models with the added complexity of covariates cannot accommodate more parameters relating to class membership. Second, findings related to the specification of inter-individual difference in trajectories were somewhat more equivocal than in the unconditional model. Unlike in the unconditional analysis, the BIC did not favor models with exclusively categorical differences between individuals in growth parameters (the LLPA and SPGM) over those with continuous inter-individual variation in growth parameters (the latent curve and latent basis models). In fact, the latent basis model fit the data better than any of models allowing exclusively categorical differences in growth parameters, including the best-fitting LLPA. However, as in the unconditional analysis, models allowing both continuous and categorical variation between individual growth parameters showed the best balance of fit and parsimony. As before, the best-fitting model in the data was a LLPA with random effects; this time a seven-class solution was favored by the BIC. Thus, the major trends of the unconditional analyses were basically replicated when covariates were added: models allowing both continuous and categorical variation in growth parameters (GMM and LLPA-RE) showed the best fit to the data, as did models in which the functional form of growth was not pre-specified (latent basis model, LLPA, and LLPA-RE).

3.2 Whole-sample predicted trajectories

Figure 1 shows the predicted trajectories at the whole-sample level for each of the best-fitting solutions for all models considered in the analysis without covariates; thickness and transparency of each line is used to indicate the proportion of individuals in the corresponding class. In both models the latent curve and latent basis models, only the mean of the intercept coefficient was significant, indicating that only initial level of depression was significantly different from zero. However, the variance of all of the growth coefficients – intercept, linear slope, and quadratic slope in the latent curve model; intercept and basis coefficient in the latent basis model – were significant. These combined findings – no growth coefficient having a significant mean, but each having a significant variance – suggest that while at the aggregate level there is no overall change in depression over time, individuals differ in the extent to which their self-rated depression changes over time.

One trend that characterized all of the mixture models is that they all placed the majority of the sample in classes defined by high CES-D scores that do not change much over time. Where the models differ is in their representation of the rest of the sample. In the SPGM, classes were differentiated from one another most by overall level; only three classes – classes 2, 4, and 6, comprising a total of 5.6% of the sample – were characterized by a significant quadratic trend, and only class 7, comprising roughly 8.0% of the sample, showed a linear increase in CES-D score over time. In the GMM, three classes – classes 2, 3, and 4 – were defined by any change over time; the remainder of the sample fell into classes 1 or 5, which where characterized by either low or high overall CES-D scores, neither of which increased or decreased over time. Thus, though there were fewer classes in the GMM than in the SPGM, roughly the same proportion of the sample, 12.1%, fell into classes characterized by any sort of change in depression over time.

The LLPA models, also placed the majority of the sample in classes defined by high CES-D scores that did not change over time; however, the nature of change in the low-CES-D classes is somewhat different from the other models. In the LLPA, roughly 20% of the sample fell outside of the three classes with high overall CES-D and no change over time, classes 6, 11, and 12. Each of these classes is characterized by a unique pattern of change that, by graphical examination, appear to represent a variety of potential functional forms; classes' trajectories range from those with small periodic fluctuations, such as classes 2 and 3, to those with large increases or decreases at one age, such as classes 1, 5, and 10. The LLPA-RE is similar in that roughly 25% of the sample fell into classes characterized by change over time, and these changes were not well-defined by any single functional form, but there were two main differences. First, here only one class, class 11, was needed to characterize individuals with high CES-D scores that did not change over time, as opposed to three. Second, the remaining classes are distinguished from one another less by differences in overall level of CES-D, and more by differences in shape, as represented graphically by the fact that trajectories overlap a great deal in overall level of CES-D scores but show wider differences in their profiles of change over time.

Results in the conditional model are presented in Figure 2. As before, neither the latent curve nor the latent basis models were characterized by a great deal of change over time. While fewer classes were found in almost all mixture models, the shapes of the trajectories defining each class were similar, and two main trends in the unconditional analysis were replicated. First, most of the sample fell into trajectories with high CES-D scores that did not change over time, with more sparsely-populated classes representing individuals with lower scores who did change over time. Second, in models in which some within-class variation in overall level of CES-D score was allowed (the GMM and the LLPA-RE), classes were less differentiated from one another by level and more by the overall shape of each trajectory.

3.3 Effects of covariates

The effects of gender, general health, college attendance by age 23, and parent-rated general health at age 13, were all tested in every model. In the latent curve model, the effect of gender on intercept was significant, $z = 6.542$, $p < .001$. No other effect on the intercept, slope, or quadratic component was significant. Results in the latent basis model were similar, with a highly significant effect of gender on the intercept (z = 7.054, p<.001) and no other effect on either the intercept or latent basis coefficient approaching significance.

Table 6 shows the results of logistic regressions using all of the covariates as predictors of class membership for each mixture model. In all models, the modal class – which in each solution was characterized by high CES-D scores and minimal change over time – is used as the reference class.

In the SPGM, members of classes 1 and 3, which showed low overall CES-D score, were more likely to be female, whereas members of class 6, which showed high overall CES-D scores, were less likely to be female. Members of class 6, the only class with higher overall CES-D score than the reference class, were also less likely to be white. Members of class 4, which was characterized by lower scores (corresponding to increased depression) at later ages, were less likely to have completed college.

In the GMM, the only covariates with any significant partial effects were gender and college. Members of classes 1 and 3, which were characterized by either lower overall CES-D score (class 1) or low CES-D score which increased over time (class 3), were more likely than members of the reference class (class 4) to be female. Members of class 1 and class 2, which was characterized by decreasing CES-D score, were also less likely to attend college than members of the reference class.

Results were relatively similar in the models without a pre-defined functional form, the LLPA and LLPA-RE, in that gender and college attendance were the only significant predictors of class membership. In the LLPA, members of classes 1 and 3, both of which are characterized by low CES-D scores which fluctuate over time, were more likely to be female than members of the reference class. By contrast, members of class 6, the only class with higher and more stable CES-D scores than the reference class, were less likely to be female; they also showed better overall parent-rated general health at the trend level.

Interestingly, members of class 4, which was defined by the decreasing CES-D scores corresponding to an increase in depression at later ages, were less likely to have attended college by age 23, but this effect was not significant for any other classes.

In the LLPA-RE, only class 3 was distinguished from the reference class by gender; thus, only members of the class with lowest and most widely-varying CES-D scores were more likely to be female. Members of this class were also less likely to have attended college by age 23, as well as members of class 1, which was also characterized by low and widely-fluctuating CES-D score, and class 6, in which CES-D score started out high but dropped precipitously at the later ages.

3.4 Individual fit

Root mean squared residual (RMSR) scores were calculated for each individual under each model using posterior-predicted individual trajectories (given Equation 14) for both the conditional and unconditional analysis; $RMSR_{Mi}$ values for each individual under each model were calculated using Equation 13 for the conditional analysis only. Their relationship to a number of covariates, including the four that were examined in the conditional model – gender, race, college attendance, and parent-rated general health – as well as the number of missing CES-D scores, were then examined. The intent with this analysis was not hypothesis-driven, but instead focused on whether, both in the unconditional analysis in which none of these covariates was included and in the conditional analysis in which the covariates were included, any given model fit individuals with certain values of the covariates better or worse.

The means of $RMSR_{Ci}$ scores for the unconditional analysis are presented in Table 7. The average $RMSR_{Ci}$ scores were relatively similar in most models but were considerably lower in the LLPA-RE than in the other models, indicating that trajectories fit more closely to individual data in the LLPA-RE than in other models. $RMSR_{Ci}$ scores were strongly related to race, general health, and college attendance, with individual fit being the strongest in white subjects, subjects with better overall general health, and subjects who attended college. Interestingly, gender was not significantly related to individual fit in any model except for the LLPA-RE, in which being female was associated with worse overall fit. Another interesting finding is that the number of missing time points was significantly related to individual fit in all models except for the LLPA and the LLPA-RE models, potentially suggesting closer fit of these models to subjects with higher degrees of missing data.

The means of $RMSR_{Ci}$ scores for the conditional analysis are shown in Table 8. As in the unconditional analysis, the LLPA-RE fits the data more closely at the individual level than any other model; however, the difference in fit was smaller than in the unconditional analysis. As before, race and college attendance were strongly related to individual fit and gender was not; however, in the conditional analysis, there was no relationship between general health and any of the models' $RMSR_{Ci}$ values except for the LLPA-RE at the trend level. The previous pattern of relationships between $RMSR_{Ci}$ and missingness were replicated: there was no significant relationship between missingness and individual model fit in the LLPA and the LLPA-RE, but significant (or, in the case of the SPGM, marginally significant) relationships to missingness in all other models.

Finally, the means of $RMSR_{Mi}$ values in the conditional analysis are shown in Table 9. Values overall were higher than the $RMSR_{Ci}$ values in either the conditional or unconditional analyses; however, the values for both of the models allowing both continuous and categorical variation in growth paramters (GMM and LLPA-RE) were the lowest, indicating closest overall fit at the individual level in these models. As with the $RMSR_{Ci}$, there was no overall relationship between gender and $RMSR_{Mi}$ values. There was, however, a significant relationship between fit and race, general health, and college attendance in every model. There was no significant relationship between fit and missingness in any model.

In summary, the fit of individual trajectories appear to follow three general patterns. First, regardless of whether the analysis included covariates or not, LLPA-RE provided closer fit to each individual's data, on average, than any other method under consideration. Second, the inclusion of covariates in the analysis did not lead to much attenuation of the relationship between covariates and fit, as indexed either using the $RMSR_{Ci}$ based on posterior-predicted individual trajectories or the $RMSR_{Mi}$ based on prior-predicted individual trajectories. Finally, when considering fit to each individual's posterior-predicted trajectory (i.e., $RMSR_{Ci}$ in either the unconditional and conditional model, the fit of LLPA and LLPA-RE is not as greatly affected by missingness of CES-D scores than the other models.

3.5 Individual Prediction

For each individual, a prediction plot given the best fitting unconditional model – the 11-class LLPA-RE – was created. These plots are shown for six individuals in the sample, in Figures 3-5. In the first panel of each plot is the 11-class LLPA-RE solution, with the lines weighted (both using line thickness and transparency) by the subject's posterior probability of belonging to each class. From this plot one can visualize how well each class represents a given individual's model-implied trajectory. In the second panel is the individual's observed and model-implied trajectories, as well as 100 predicted trajectories generated by the bootstrap methods of Pek, Bauer, and Losardo (2011) using the covariance matrix of model parameters. By visualizing the uncertainty around the individual predicted trajectory nonparametrically, one can get a sense of the error around each individual observation.

In Figure 3, the plots for subjects 1266 and 1401 are shown; these two individuals had $RMSR_{Ci}$ scores at the median for the sample, and thus they represent the exact midpoint of individual model fit. In Figure 4, the plots for subjects 286 and 560 are shown; these two individuals had $RMSR_{Ci}$ scores at the 75th percentile and thus represent relatively poor individual model fit. In Figure 5, the plots for subjects 550 and 691 are shown; these two individuals had $RMSR_{Ci}$ scores at the 25th percentile and thus represent relatively close individual model fit.

Finally, in order to demonstrate the graphical capabilities of LLPA with covariates, predicted trajectories were created by using hypothetical configurations of $\boldsymbol{z}_i$ in Equation 5 to create prior probabilities of class membership. Figure 6 shows one such plot, which compares subjects on the basis of gender and college attendance (holding race constant at white, and general health constant at the sample average). The plot provides insight into one trend that might not have been apparent from examination of the model parameters alone: while female college attendees have higher CES-D scores than non-attendees, indicating generally lower depression for the entire study period, male college attendees and non-attendees seem to diverge in depression only in the later ages.

The current analyses compare the LLPA framework to other analytic approaches in the context of only one dataset. It may be the case that LLPA was favored by fit indices due to idiosyncratic characteristics of the NLSY data. This possibility is explored next using a brief proof-of-concept simulation, which examines the sensitivity of a number of these models to random noise in the data.

3.6 Proof-of-concept simulation

In the preceding analyses, both variations of the LLPA framework – with and without a random effect of level – were favored by the BIC over the latent basis and latent curve models (as well as other models). At the same time, all models tested gave the impression of minimal change over time for most individuals. This conclusion stems from the lack of significant growth coefficients in either the latent curve or latent basis models, as well as as the fact that in the mixture models most individuals were in groups characterized by minimal change. These general findings, in combination, raise the concern that the impression of LLPA models' superior fit may in fact be driven by overfitting in the absence of systematic change in the data.

In order to address these concerns, I simulated data from a random-intercept model with no growth of any sort and tested whether an LLPA would be spuriously favored over an intercept-only latent curve model. In order to do this, I simulated six time points for 1686 cases (as there were 1686 individuals in the original sample). The random intercept was set to have a mean of 15.24 and a variance of 3.35, as these were the mean and variance of the random intercept in the latent curve model fit in the unconditional analysis. Similarly, residual variance at each time point was set at 3.166, as this was the estimated residual variance in this latent curve model. The slope coefficient was fixed at zero and had zero variance.

While generating data from a latent curve model may partially address the concern that an LLPA might pick up random noise in the data and be spuriously chosen over a LCM, the original data is ”noisy” in two ways that simulated data typically is not. First, where the simulated data is fully continuous, the original data was on a 20-point scale and scores could only take on integer values. Second, there was no missingness in the simulated data, whereas there was a moderate degree of missingness in the original data. To rectify these two discrepancies, I altered the simulated dataset such that all values were rounded to the nearest whole number, and values were deleted to simulate the pattern of missingness shown in Table 2.

The models tested were an intercept-only LCM, as well as an LCM with intercept and linear slope components; additionally, LLPA models with and without random effects for levels were fit to the data. BICs for all models tested are shown in Table 10, ranked from best to worst values. As predicted, an intercept-only LCM with equal time variances was the best-fitting model. An LCM with both intercept and linear slope was attempted, but yielded a Hessian matrix that was not positive definite. Among the LLPA models, the BIC favored a one-class model when a random effect for level was included, and a five-class model when the random effect was excluded. However, regardless of the fact that an LLPA without a random effect spuriously favored a solution with more than one class, the latent curve model was nevertheless favored over any of the LLPA models.

Model-implied trajectories for the 5-class LLPA solution shown in the second dataset are shown in Figure 7. Rather than several nonparametric shapes of change, as was observed in the original analysis (shown in the middle panel of Figure 1), the model-implied trajectories here are simply five straight lines, with no change across time and differences exclusively in overall level. Thus, the LLPA does not, in this case, appear to spuriously find systematic change in the data.

4 DISCUSSION

This thesis introduced longitudinal latent profile analysis (LLPA), a mixture model which allows for maximum flexibility in the shape of intra-individual change and inter-individual differences in longitudinal analyses. The LLPA models growth in terms of two parameters – level and shape – and allows these parameters to vary categorically between people. This model was extended to allow for a random effect of level within each class (LLPA-RE), which allows class membership to reflect more shape-based differences across individuals. The LLPA-based methods were applied to a study of depression in a nationally representative sample of adolescents, and the fit of these models was compared with a number of comparable latent curve models and parametric mixture models at the whole-sample level using the Bayesian Information Criterion (BIC). This thesis also examined the use of individual fit statistics, particularly the individual root mean squared residual (RMSR), to examine the different models' fit at the individual level, and examined graphical methods for individual-level inference.

4.1 Fit at the whole-sample level

All models tested, including the LLPA and the LLPA-RE, found that trajectories with high overall CES-D scores and virtually no change in scores over time applied to the bulk of the sample. These scores indicate that most of the sample was characterized by a lack of depressive symptomatology that was stable over time. The LLPA-based methods, however, also disaggregated the sample at the lower levels of CES-D scores, breaking a minority of the sample into smaller groups characterized by high levels of depression, which showed various patterns of change over time.

The main question regarding these findings is whether the non-parametric shapes uncovered by LLPA and LLPA-RE are valid points from which to draw inference, or simply capturing random noise in the data. A number of the current findings offer a good deal of evidence that the LLPA analyses picked up signal, and not noise, in the data. The BIC was reliably lowest in the LLPA-RE, both in models with and without covariates, and importantly the BIC reflects both parsimony and fit. This finding likely reflects the minimally restrictive nature of the LLPA-RE, which neither forces all individuals to have the same shape of growth, nor the same overall level of depression as everyone in their class. Interestingly, after the LLPA-RE the second-lowest BIC scores tended to be observed in the GMM, suggesting the possibility that allowing flexibility in the representation of between-person differences in growth parameters may be the best way to optimize fit and parsimony simultaneously.

Further evidence for the LLPA's lack of susceptibility to conflating random noise with systematic change comes from the proof-of-concept simulation. These results suggest that, in the presence of only level differences, neither the LLA nor its random-level extension will erroneously find nonlinear shapes in the data. The LLPA-RE only found one class, and this class was only characterized by a significant mean and variance for level, with no change in shape. Even an LLPA solution which erroneously divided the sample into classes found these classes to be differentiated only by level, with complete stability of scores across times for all groups. More systematic simulations in future work will further assess LLPA methods' sensitivity to error variance in the lack of any meaningful change in the data. One important focus of potential future work will be investigating the effect of heteroscedasticity of error – importantly, the simulation assumed that error variance was distributed identically for all classes, but this may not be a tenable assumption in the CES-D.

The question of error heteroscedasticity in the CES-D presents one obvious argument for LLPA's findings at the lower end of the CES-D's range being attributable to random noise as opposed to systematic change: responses at the extreme end of the scale are necessarily more error-prone than observations closer to the mean. While the CES-D has been shown to have comparable internal consistency and test-retest reliability in both depressed and non-depressed populations (Radloff, 1997), it is important to consider the possibility that non-parametric shapes among those with higher overall degrees of depression are an artifact of random, unsystematic fluctuation in depression.

It is also important to consider, however, the difference between unsystematic error and meaningful instability across time – i.e., state variance as opposed to trait variance in depression. This distinction has been raised with regard to the CES-D in adolescents by Dumenci and Windle (1996), who applied a latent trait-state (LTS) model. Using the full CES-D, the authors determined that there was a significant component relating CES-D scores to trait-level depression, as well as a significant amoung of state-level variance. They hypothesize that the relatively low test-retest reliability of the CES-D (Lin and Ensel, 1984) may in fact be a byproduct of the fact that the CES-D measures a great deal of state variance in depression, which, though it represents meaningful time-to-time fluctuation, is often conflated with measurement error (Nesselroade, 1988).

In essence, the LLPA conducted here can potentially be seen as a different way of casting questions of state vs. trait measurement of depression: whereas previous work has isolated specific variables (i.e., CES-D items) that are more related to either state or trait depression, the current LLPA may capture differences between specific *cases* in terms of state- and trait-level depression. It may be of clinical and scientific significance to distinguish cases with ”reactive” depressive symptomatology – i.e., individuals who become more depressed following a life event – from those with more constant depressive symptoms, and to determine possible connections to these disparate experiences (Hankin, Mermelstein, and Roesch, 2007). Importantly, the use of LLPA to distinguish such cases does not necessitate interpreting each predicted increase or decrease in CES-D score for each class; it may be of interest to simply differentiate individuals who experience change from those who do not. By contrast, parametric mixtures such as SPGM, in hypothesizing a particular form of growth (e.g., linear or quadratic), are much more commonly directly interpreted as being representative of a particular type of growth or decay.

4.2 The effect of covariates

The effect of covariates on growth was relatively consistent across models. All found that female participants generally reported higher overall levels of depressive symptomatology, which is consistent with the widely-reported finding that women experience depression with higher frequency than men (e.g. Nolan-Hoeksema, 1990; Kessler et al., 2005). Results related to race and parent-reported general health were generally minimal, with only a few inconsistent findings relating CES-D score to either covariate.

The results present a complicated picture of the relationship between college attendance and CES-D score, with only the mixture models detecting any difference in probability of college attendance between the classes. However, all of the mixture models isolated at least one class with a lower probability of college attendance than the reference class; in the SPGM, GMM, and LLPA, this was a class with decreasing values of CES-D, corresponding to increased depressive symptomatology at the later ages. Given that educational attainment and SES are highly correlated, and that low SES is strongly linked to increased depression (Lorant et al., 2003; Everson, Maty, Lynch, and Kaplan, 2002), it may be the case depressive symptomatology worsens at later ages among college non-attendees after the socioeconomic effects of not having attended college (e.g., instability of employment) begin to manifest.

Interestingly, the LLPA-RE found more systematic class differences in college attendance, including a decreased probability of college attendance among members of three classes, each of which were characterized by low CES-D score at some point in development. This finding stands somewhat in contrast to the results of the SPGM, GMM, and LLPA, that individuals with increasing levels of depression across early adulthood, but not those with generally high levels of depression or decreasing levels of depression, are less likely to go to college. This discrepancy raises the possibility that more strictly parameterized methods (e.g., SPGM, GMM, and LLPA without random effects for level) may find that a covariate is linked to a given shape which no longer holds once a more flexible model is applied.

4.3 Individual-level fit and prediction

A number of the interesting features of LLPA are illuminated when considering the model in terms of inference at the individual level. In particular, the individual RMSR as presented by Coffman and Milsap (2006) offers two different types of conclusions about the relative fit of all models considered. The $RMSR_{Ci}$, which compares each individual's observed trajectory to the one that would be predicted from that individual's posterior probability of belonging to each class, indicates that in general the LLPA-RE yields predicted trajectories for each individual that are closest to his or her observed data. When covariates are included in the model, the LLPA-RE is still the closest fit to the data; however this difference in fit is somewhat attenuated. The second individual fit statistic considered, $RMSR_{Mi}$, represents something different from the $RMSR_{Ci}$: in conditioning on covariate-based prior probabilities, it indexes how closely the covariates predict the observed trajectory for each individual in the sample. Despite the fact that there were relatively few differences between any of the models in findings at the aggregate level with respect to covariates, both the LLPA-RE and the GMM had considerably lower values of $RMSR_{Mi}$ than any of the other models, suggesting the possibility that models allowing the most flexible representation of inter-individual differences will also provide the closest covariate-based prediction.

Comparing the individual RMSR scores – both $RMSR_{Ci}$ and $RMSR_{Mi}$ – to a number of covariates can theoretically be an informative way to differentiate what features characterize individuals with particularly close or less close fit to the model. However, in the current analysis, all RMSR values were generally correlated with general health, college attendance, and race, with models fitting more closely to white college attendees in better overall health. The discrepancy between these individual-level results and the aggregate-level covariate effects – i.e., the fact that RMSR values were related to these covariates, even in the absence of these covariates having an effect in the model – is likely due to the fact that covariate effects in the aggregate model were partial statistics, whereas zero-order correlations were examined here. Another possibility for this discrepancy is the difference in the relationship that the RMSR and covariates have to the time-varying indicators. RMSR is a measure of variability, but covariates predict mean levels of the time-varying indicator – if mean levels of an indicator vary little according to the covariates but certain values of the covariates are associated with greater variance of the indicator, values of the RMSR may still be associated with that covariate.

Interestingly, however, values of both $RMSR_{Ci}$ and $RMSR_{Mi}$ were correlated with the number of missing CES-D scores in all of the models except for the LLPA and LLPA-RE. This may suggest these models may be particularly helpful in fitting to individual data under conditions of missingness; conversely, it may suggest that LLPA-based models overfit to whatever data points are present as opposed to obtaining a maximally generalizable solution in the presence of missing data.

Whether LLPA-RE's close fit to individual data is a strength or a weakness is an open question. As with findings at the aggregate level, concern over whether these models pick up on signal vs. noise is at the crux of this question. However, unlike the measures of interest in the whole-sample analysis (predominantly the BIC), the measure of interest here, RMSR, does not incorporate measures of model parsimony. Thus there is less conclusive evidence within the scope of this analysis to inform questions of individual level findings' generalizability, and further examination is required.

However, regardless of the ambiguity over whether LLPA-RE's close fit to the data is generally a strength or a weakness, in fitting more closely to individual data points, LLPA-RE can at least be said to incorporate more information about each individual's path into the final analysis. Thus, it may be that LLPA-RE is more consistent with the goal of building psychological science from the ”bottom-up” (Molenaar and Campbell, 2009). Whereas psychological research has tended to be guided by the goal of finding trends that apply to everyone in a given population, there are many processes for which whole-sample and individual-level patterns of change are not able to be equated with one another (i.e., non-ergodic processes; Molenaar, 2004). Perhaps, thus, a method such as LLPA-RE, which allows a great deal of individual variation to be taken into account in the fitting of the model for the whole sample, represents a good option for examining data under conditions in which individual and whole-sample level trajectories cannot be equated.

The examination of model-predicted trajectories for each individual can provide some insight into these processes at the individual level. Given that LLPA and LLPA-RE both allow individuals to vary according to a greater number of shapes than a parametric mixture model or a latent curve model, individual-level prediction may be more interesting and informative in the LLPA framework.

As with many other individual-level analyses, it is important to consider how to make a scientifically useful inference based on predictions for individuals, each of whom represents a very small fraction (in the current analysis, 1/1686) of the sample, which itself is a small fraction of the population (e.g., Castro-Schilo and Ferrer, 2013). In the current analysis, individuals were chosen for examination based on individual-level fit; however, individuals can be chosen for examination as representative cases on the basis of whatever is of interest to the researcher. In some cases, a relatively rare condition – for instance, hard drug use in a sample of young adolescents – may be of interest to the researcher; thus one might want to plot the individual predicted plots of the few individuals who meet that condition, and compare them to the overall sample mean trajectory. Additionally, it may be of interest to examine the model-predicted trajectories of cases who are known to exert a high degree of influence on the overall model (Pek and MacCallum, 2010; Sterba and Pek, 2012), in order to determine what sorts of trajectories particularly influential cases tend to follow.

Furthermore, predicted individual trajectories based on prior probabilities of class membership given values of the covariates may be of particular interest. Methods for making inferences about predicted trajectories based on covariate values have been developed in the latent curve modeling literature (Preacher, Curran, and Bauer, 2006; Curran, Bauer, and Willoughby, 2006), and it could be particularly interesting to compare predicted trajectories for hypothetical individuals in LLPA based on a vector of covariates. The use of bootstrapping can approximate confidence intervals around predicted trajectories to help determine the point at which individuals with different values of covariates become significantly different from one another; this can help in understanding the complicated interplay between covariates and the outcome of interest over time.

4.4 Limitations and future directions

The current thesis introduces LLPA and LLPA-RE and compares these models to a number of widely-used models only in their fit to one dataset. Specifically, the current dataset represents a large, diverse group of individuals (N=1686) measured using an instrument with relatively coarse gradations (with possible CES-D scores ranging from 0 to 20). It will thus be of interest to examine the relative performance of all of these models when either sample size is small, measurement is more precise, or both. Furthermore, in the current study there were relatively few significant covariate effects in any model considered. Thus, the relative power of the LLPA-based methods to detect meaningful between-person differences based on values of covariates remains to be seen.

A more systematic examination of the issues facing the interpretation of LLPA could be answered using a simulation study. However, it is difficult to establish a valid simulation condition that meaningfully examines the differences between LLPA, LLPA-RE, and comparable models. If one were to simulate data in which a mixture of trajectories existed, mixture models would almost definitely fit better than models without a mixture component; similarly, if one simulated data in which observations followed non-parametric growth patterns, models which did not impose a functional form would necessarily fit better.

What may be of interest is to compare LLPA to a variable-based method of examining trajectories according to multiple freely-estimated growth factors. In particular, it may be interesting to examine LLPA in relation to Tuckerized curves (Tucker, 1958), or exploratory latent growth curve models, a new extension of Tuckerized curves in the SEM framework (Grimm, Steele, Ram, and Nesselroade, 2013). In disaggregating the overall trajectory into multiple, non-parametric growth curves, this method allows for a similarly flexible representation of growth to LLPA: whereas these methods decompose variability into multiple latent variables, LLPA decomposes variability into cases. Comparing these two approaches may be particularly interesting given the equivalency between a K-class model and a K+1-variable factor analysis (Bartholomew and Knott, 1999). However, this equivalency does not hold for higher-order moments, which also factor into the overall likelihood of the data (Bauer and Curran, 2004). Thus, it may be interesting to alter the skewness and kurtosis the multivariate distribution, and see under which other conditions – e.g., sample size, measurement, distribution of the dependent variable – each of these models provides a better fit to the data.

Despite these limitations and need for future research, the current thesis represents the introduction of a model which may be particularly useful for making both whole-sample and individual-level inferences in longitudinal data. By disaggregating the indicators of a latent profile analyis into time-invariant level and time-varying shape, the LLPA obtains the most flexible representation of inter- and intra-individual difference. Further, the random-level extension of LLPA, the LLPA-RE, allows even more flexibility in inter-individual differences. These data were applied to an empirical dataset, and assessments of fit at the aggregate and individual levels demonstrated the advantage of LLPA and LLPA-RE over comparable models. By examining the relative strengths and weaknesses of these methods systematically, we can establish guidelines by which researchers can choose the best method for their data, sample, and research goals.

5 FIGURES AND TABLES

Table 1: Summary of models under consideration

Functional Form of Growth	**Treatment of Growth Parameters**		
	Continuous	Continuous + Categorical	Categorical
Specified	Standard latent curve model (LCM)	Growth mixture model (GMM)	Semiparametric growth model (SPGM)
Unspecified	Latent basis model	Random-effects LLPA	Longitudinal latent profile analysis (LLPA)

Table 2: Percent CES-D scores present at each time point

	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6
Time 1	**95.8**					
Time 2	92.1	**95.0**				
Time 3	86.7	86.3	**89.9**			
Time 4	84.5	84.3	80.8	**88.4**		
Time 5	84.1	83.9	79.8	81.2	**88.0**	
Time 6	83.6	83.1	79.2	79.2	80.7	**87.3**

*N*ote. * indicates $p < .05$; ** indicates $p < .01$; ^{t} indicates $p < .10$.

Table 3: Frequency of each response at all time points, in percentage points

	1	2	3	4
How often has respondent felt depressed in the past month				
	1	2	3	4
Time 1	3.1	7.9	53.0	36.1
Time 2	3.4	11.8	59.0	25.9
Time 3	2.9	8.2	49.5	39.4
Time 4	2.5	7.4	49.6	40.4
Time 5	3.5	10.0	52.3	34.1
Time 6	2.3	6.8	48.0	42.9
How often has respondent been a happy person in the past month				
	1	2	3	4
Time 1	10.1	48.2	36.5	5.2
Time 2	9.0	44.7	40.6	5.6
Time 3	8.0	50.8	35.7	5.4
Time 4	7.6	51.2	37.7	3.6
Time 5	6.6	48.7	40.4	4.2
Time 6	8.6	50.7	36.2	4.5
How often has respondent felt down or blue in the past month				
	1	2	3	4
Time 1	3.4	10.8	55.3	30.5
Time 2	3.3	11.0	58.4	27.3
Time 3	2.1	8.2	55.6	34.1
Time 4	1.6	7.9	54.7	35.8
Time 5	2.3	9.4	54.9	33.4
Time 6	1.9	7.0	51.3	39.9
How often has respondent felt calm or peaceful in the past month				
	1	2	3	4
Time 1	15.0	54.0	28.2	2.8
Time 2	13.5	53.5	30.5	2.5
Time 3	13.1	55.9	28.4	2.6
Time 4	11.9	58.3	27.5	2.3
Time 5	10.2	56.3	31.2	2.3
Time 6	12.5	56.7	29.0	1.8
How often has respondent been a nervous person in the past month				
	1	2	3	4
Time 1	2.0	6.1	26.2	65.7
Time 2	2.0	4.7	30.2	63.1
Time 3	0.7	4.4	25.5	69.4
Time 4	1.0	3.5	23.9	71.5
Time 5	1.3	3.9	24.8	70.0
Time 6	0.7	3.4	23.7	72.1

Table 4: Comparative fit of all unconditional models

Latent Curve

–	BIC	Rank
Equal var.	**39776.508**	**42**
Non-equal var.	39793.388	43

Latent Basis

–	BIC	Rank
Equal var.	**39766.271**	**40**
Non-equal var.	39776.027	41

SPGM

Num. Classes	BIC	Rank
2	40432.253	50
3	39971.609	47
4	39836.793	44
5	39750.211	38
6	39705.782	35
7	39671.196	30
8	39662.26	29
9	39662.218	28
10	39661.706	27
11	**39660.285**	**25**
12	–	
13	–	

LLPA

Num. Classes	BIC	Rank
2	40411.58	49
3	39994.823	48
4	39841.244	46
5	39763.247	39
6	39729.795	37
7	39692.754	33
8	39661.558	26
9	39642.824	23
10	39626.002	21
11	39615.559	20
12	**39604.157**	**16**
13	39609.26	17

GMM

Num. Classes	BIC	Rank
2	39717.107	36
3	39613.785	19
4	39559.662	11
5	**39551.931**	**10**
6	39563.825	12
7	39577.347	13
8	39595.423	15
9	39612.414	18
10	39640.866	22
11	39656.083	24
12	39689.338	32
13	39705.761	34

LLPA with Freed Level

Num. Classes	BIC	Rank
2	39678.352	31
3	39588.703	14
4	39841.244	46
5	39470.409	9
6	39433.466	8
7	39406.201	7
8	39389.746	6
9	39374.63	3
10	39370.838	2
11	**39359.839**	**1**
12	39375.843	4
13	39384.591	5

*N*ote. ”Equal variance” refers to homogenous residual variances for all time points; ”non-equal variance” refers to heterogeneous residual variances across time points.

Table 5: Comparative fit for all conditional models

Latent Curve

–	BIC	Rank
Equal var.	**35178.851**	**18**
No equal var.	35195.42	23

SPGM

Num. Classes	BIC	Rank
2	35705.636	30
3	35351.405	27
4	35267.417	25
5	35189.623	20
6	**35155.907**	**16**
7	35156.673	17
8	–	

GMM

Num. Classes	BIC	Rank
2	35143.525	14
3	35063.002	10
4	**35002.98**	**6**
5	35023.218	7
6	35051.545	9
7	35078.534	11
8	–	

Latent Basis

–	BIC	Rank
Equal var.	**35139.109**	**13**
No equal var.	35154.36	15

LLPA

Num. Classes	BIC	Rank
2	35686	29
3	35371.517	28
4	35278.115	26
5	35208.934	24
6	**35182.172**	**19**
7	35190.611	21
8	35193.938	22

LLPA with Freed Level

Num. Classes	BIC	Rank
2	35099.074	12
3	35025.469	8
4	34993.725	5
5	34968.272	4
6	34954.296	3
7	**34947.11**	**1**
8	34947.11	2

*N*ote. ”Equal variance” refers to homogenous residual variances for all time points; ”non-equal variance” refers to heterogeneous residual variances across time points.

Table 6: Covariate effects for the conditional analyses

	Male		Gen. Health		College		White	
SPGM – Reference Class = 2								
1	$-1{\cdot}197^{**}$	(0·422)	0·069	(0·168)	$-1{\cdot}309^{t}$	(0·791)	0·217	(0·303)
3	$-0{\cdot}571^{**}$	(0·21)	−0·036	(0·112)	−0·225	(0·238)	−0·053	(0·211)
4	0·358	(0·463)	−0·004	(0·245)	$-1{\cdot}493^{*}$	(0·688)	−0·459	(0·522)
5	$-0{\cdot}724^{t}$	(0·405)	−0·357	(0·252)	−0·443	(0·394)	−0·476	(0·391)
6	$0{\cdot}705^{**}$	(0·223)	$-0{\cdot}284^{t}$	(0·166)	−0·31	(0·257)	$-0{\cdot}597^{**}$	(0·216)
GMM – Reference Class = 4								
1	$-1{\cdot}636^{**}$	(0·416)	0·23	(0·162)	$-1{\cdot}008^{*}$	(0·499)	0·291	(0·287)
2	0·146	(0·417)	0·192	(0·21)	$-1{\cdot}232^{*}$	(0·547)	−0·343	(0·49)
3	$-0{\cdot}838^{*}$	(0·384)	−0·08	(0·221)	−0·542	(0·349)	−0·184	(0·322)
LLPA – Reference Class = 5								
1	$-1{\cdot}163^{**}$	(0·388)	0·074	(0·168)	−1·274	(0·779)	0·183	(0·306)
2	$-0{\cdot}776^{t}$	(0·413)	−0·376	(0·265)	−0·44	(0·45)	−0·493	(0·417)
3	$-0{\cdot}539^{**}$	(0·208)	−0·051	(0·113)	−0·196	(0·243)	−0·04	(0·201)
4	0·182	(0·542)	0·013	(0·271)	$-1{\cdot}684^{*}$	(0·659)	−0·363	(0·496)
6	$0{\cdot}704^{**}$	(0·224)	$-0{\cdot}293^{t}$	(0·164)	−0·301	(0·249)	$-0{\cdot}585^{**}$	(0·217)
LLPA-RE – Reference Class = 7								
1	−0·523	(0·603)	0·161	(0·291)	$-1{\cdot}497^{*}$	(0·68)	0·435	(0·527)
2	$-1{\cdot}498^{t}$	(0·84)	−0·112	(0·261)	−0·791	(0·711)	−0·851	(0·836)
3	$-0{\cdot}944^{*}$	(0·368)	0·136	(0·209)	$-1{\cdot}249^{**}$	(0·426)	0·157	(0·403)
4	−0·633	(0·382)	−0·029	(0·247)	−0·249	(0·404)	−0·203	(0·394)
5	$0{\cdot}308^{t}$	(0·585)	0·364	(0·221)	$-1{\cdot}795^{t}$	(1·086)	−0·553	(0·487)
6	0·096	(0·513)	0·157	(0·283)	$-1{\cdot}306^{*}$	(0·651)	−0·81	(0·691)

*N*ote. * indicates $p < .05$; ** indicates $p < .01$; ^{t} indicates $p < .10$.

Table 7: Individual $RMSR_{Ci}$ values: Unconditional analysis

	Mean	**SD**	**Correlation coefficients of covariates**				
			Female	Genhealth	College	White	Num. Missing
Latent Curve	1.37	0.687	0.006	0.088**	-0.197**	-0.140**	0.067**
Latent Basis	1.45	0.725	0.010	0.080**	-0.197**	-0.146**	0.067**
SPGM	1.47	0.666	0.033	0.108**	-0.197**	-0.158**	0.057*
LLPA	1.39	0.561	0.034	0.094**	-0.180**	-0.149**	0.009
GMM	1.44	0.641	0.028	0.090**	-0.189**	-0.151**	0.064**
RE-LLPA	1.23	0.455	0.051*	0.067**	-0.147**	-0.125**	-0.022

*N*ote. * indicates $p<.05$; ** indicates $p<.01$; ^{t} indicates $p<.10$.

Table 8: Individual $RMSR_{Ci}$ values: Conditional analysis

	Mean	**SD**	**Correlation coefficients of covariates**				
			Female	Genhealth	College	White	Num. Missing
Latent Curve	1.409	0.698	0.010	0.090	-0.193**	-0.136**	0.057*
Latent Basis	1.420	0.705	0.007	0.084	-0.192**	-0.141**	0.053*
SPGM	1.520	0.680	0.027	0.105	-0.168**	-0.121**	0.043^t
LLPA	1.506	0.673	0.028	0.102	-0.160**	- 0.125**	0.035
GMM	1.446	0.651	0.050^t	0.080	-0.168**	- 0.131**	0.057*
RE-LLPA	1.306	0.526	0.071**	0.047^t	-0.119**	- 0.101**	0.000

*N*ote. * indicates p <.05; ** indicates p<.01; ^{t} indicates p<.10.

Table 9: Individual $RMSR_{Mi}$ values: Conditional analysis

	Mean	**SD**	**Correlation coefficients of covariates**				
			Female	Genhealth	College	White	Num. Missing
Latent Curve	2.245	1.018	-0.016	0.054*	-0.161**	-0.101**	0.041
Latent Basis	2.253	1.017	-0.023	0.052*	-0.157**	-0.106**	0.038
SPGM	2.245	1.017	-0.017	0.055*	-0.162**	-0.103**	0.041
LLPA	2.238	1.018	-0.016	0.053*	-0.160**	-0.104**	0.040
GMM	1.669	0.932	-0.003	0.074**	-0.179**	-0.116**	0.035
LLPA-RE	1.553	0.772	0.019	0.080**	-0.183**	-0.133**	0.050^t

*N*ote. * indicates $p < .05$; ** indicates $p < .01$; ^{t} indicates $p < .10$.

Table 10: Model fit for proof-of-concept simulation

Model	Num. classes	BIC
Latent curve model, intercept only, equal time variances	–	49679.473
Latent curve model, intercept and slope, equal time variances**	–	49694.045
LLPA-RE	1	49715.123
Latent curve model, intercept only, unequal time variances	–	49721.7
Latent basis model	–	49744.691
LLPA-RE	2	49750.754
LLPA-RE	3	49787.871
LLPA	5	49927.437
LLPA	6	49949.162
LLPA	4	50025.623
LLPA	3	50270.141
LLPA	2	50993.47
LLPA	1	53814.501

Note: ** Denotes that the solution is not trustworthy due to probable convergence at a local maximum.

Figure 1: Whole-sample predicted trajectories under all unconditional models

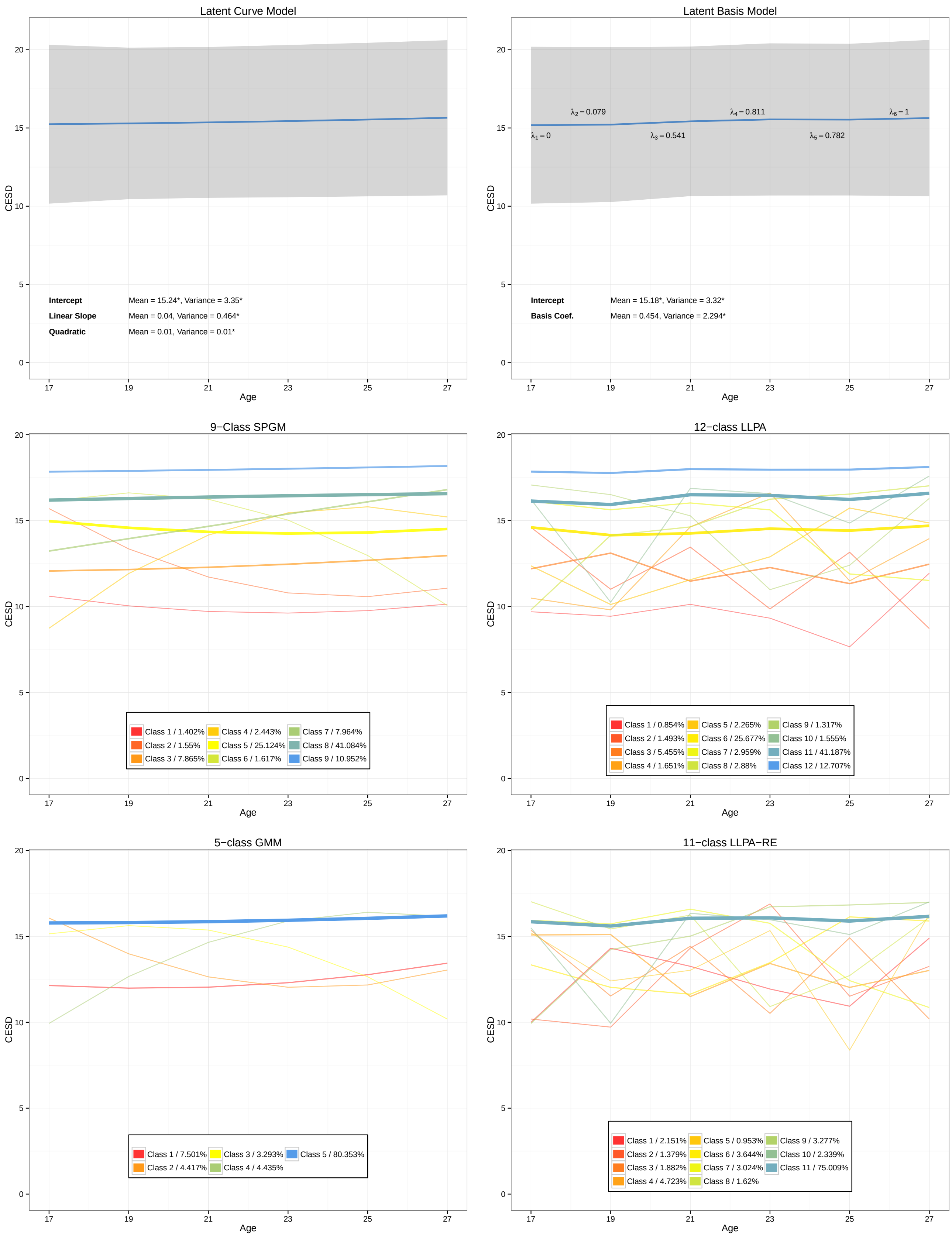


Figure 2: Whole-sample predicted trajectories under all conditional models

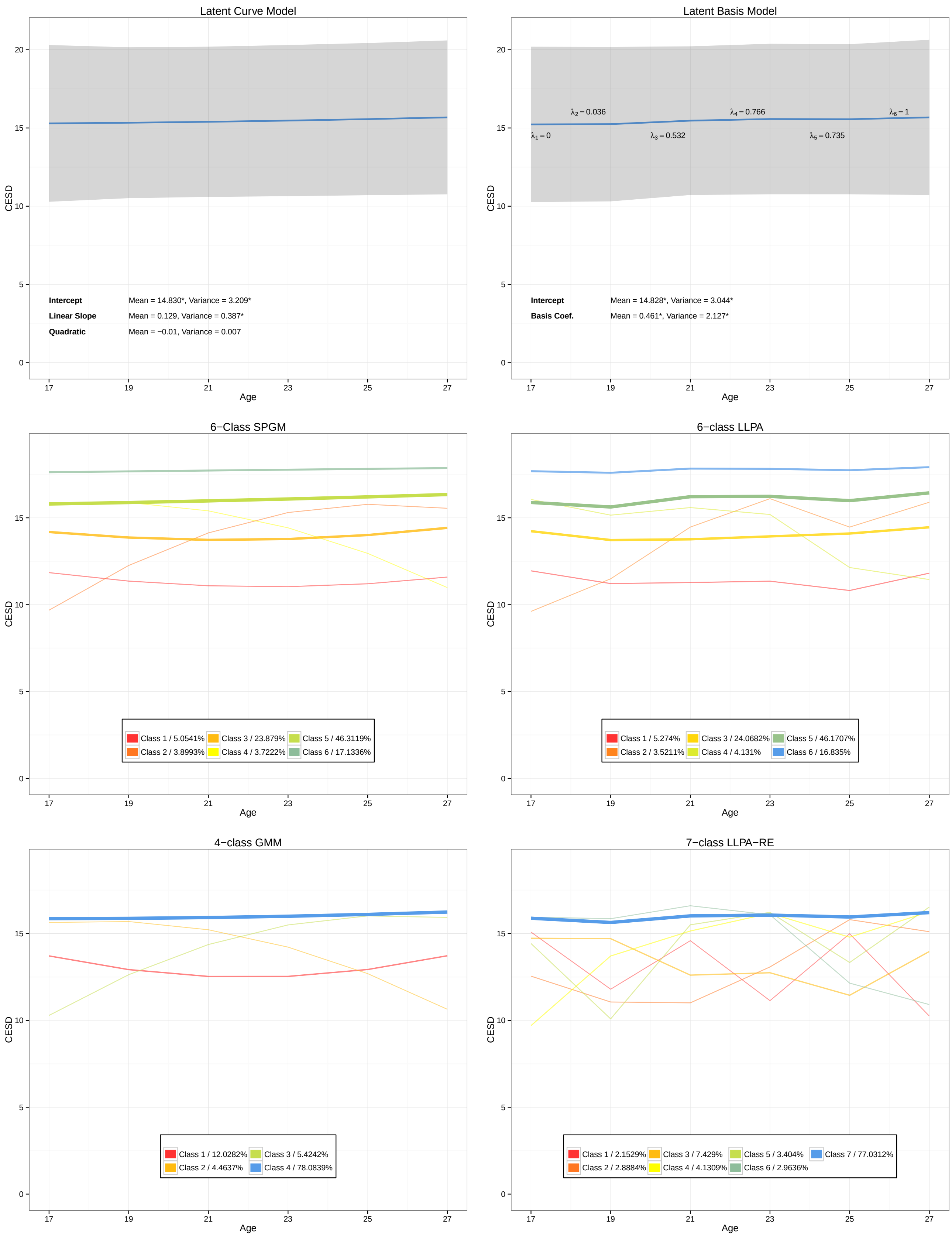


Figure 3: Predicted trajectories for subjects at the 50th percentile of RMSR

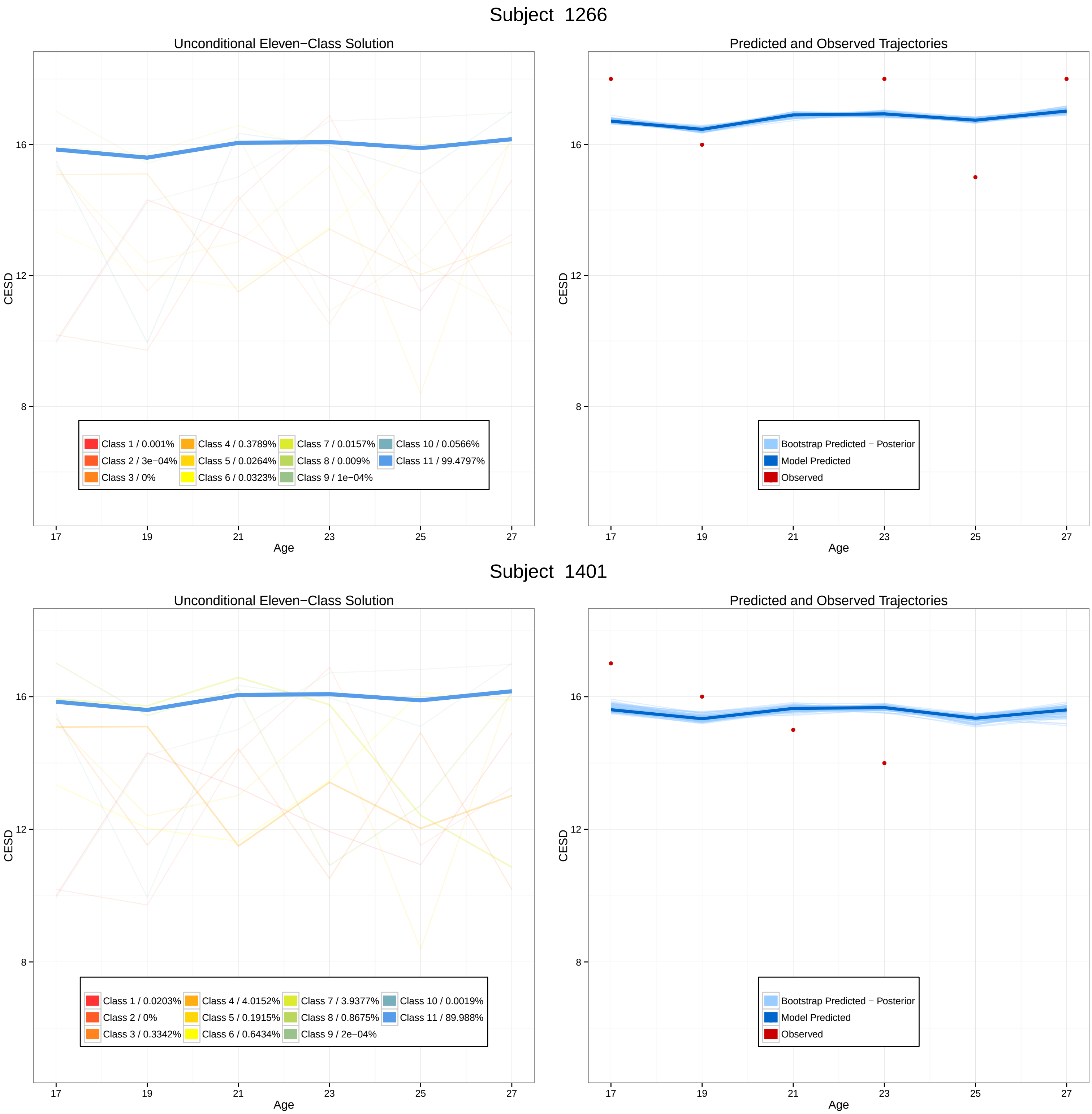


Figure 4: Predicted trajectories for subjects at the 75th percentile of RMSR

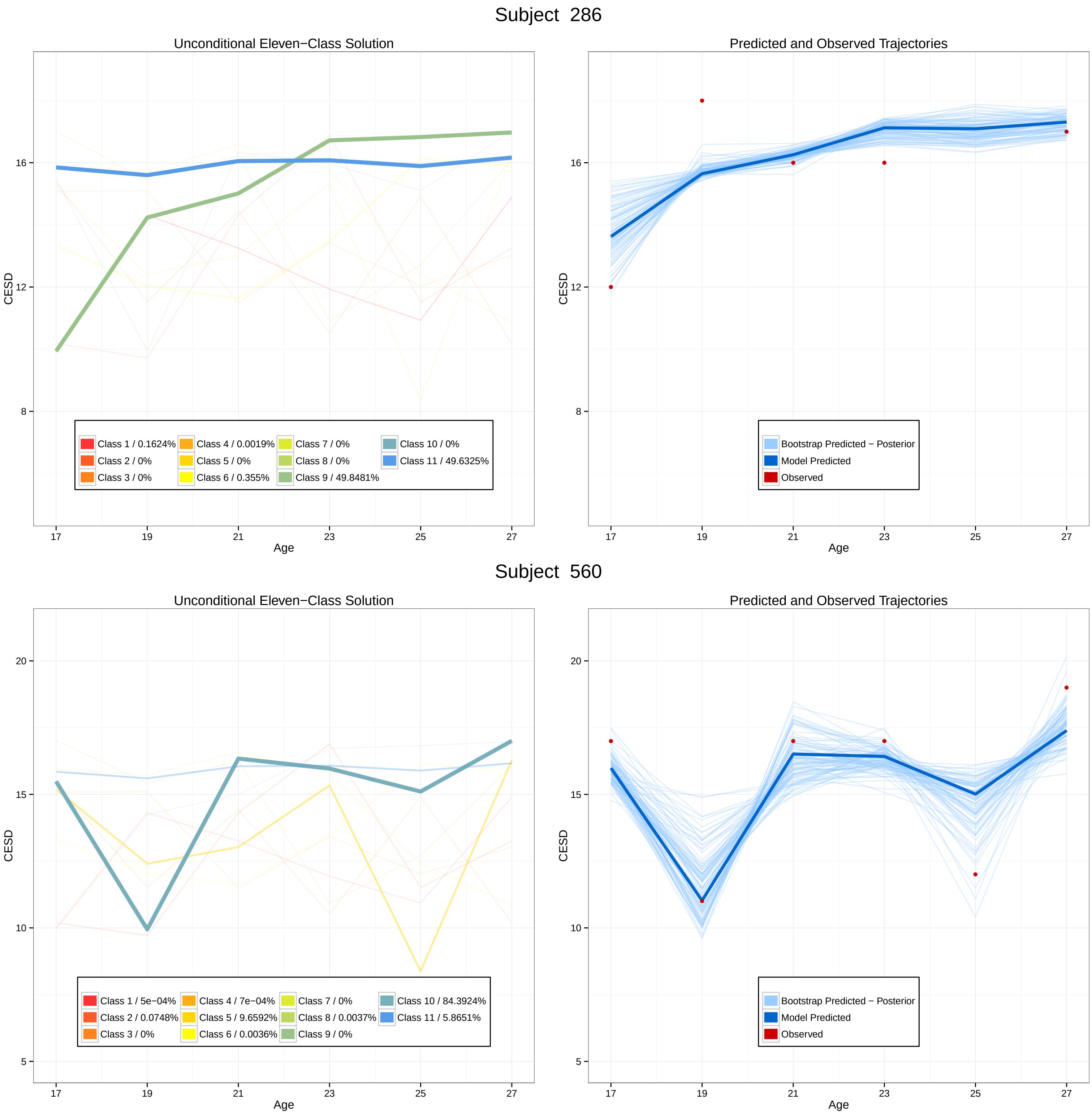


Figure 5: Predicted trajectories for subjects at the 25th percentile of RMSR

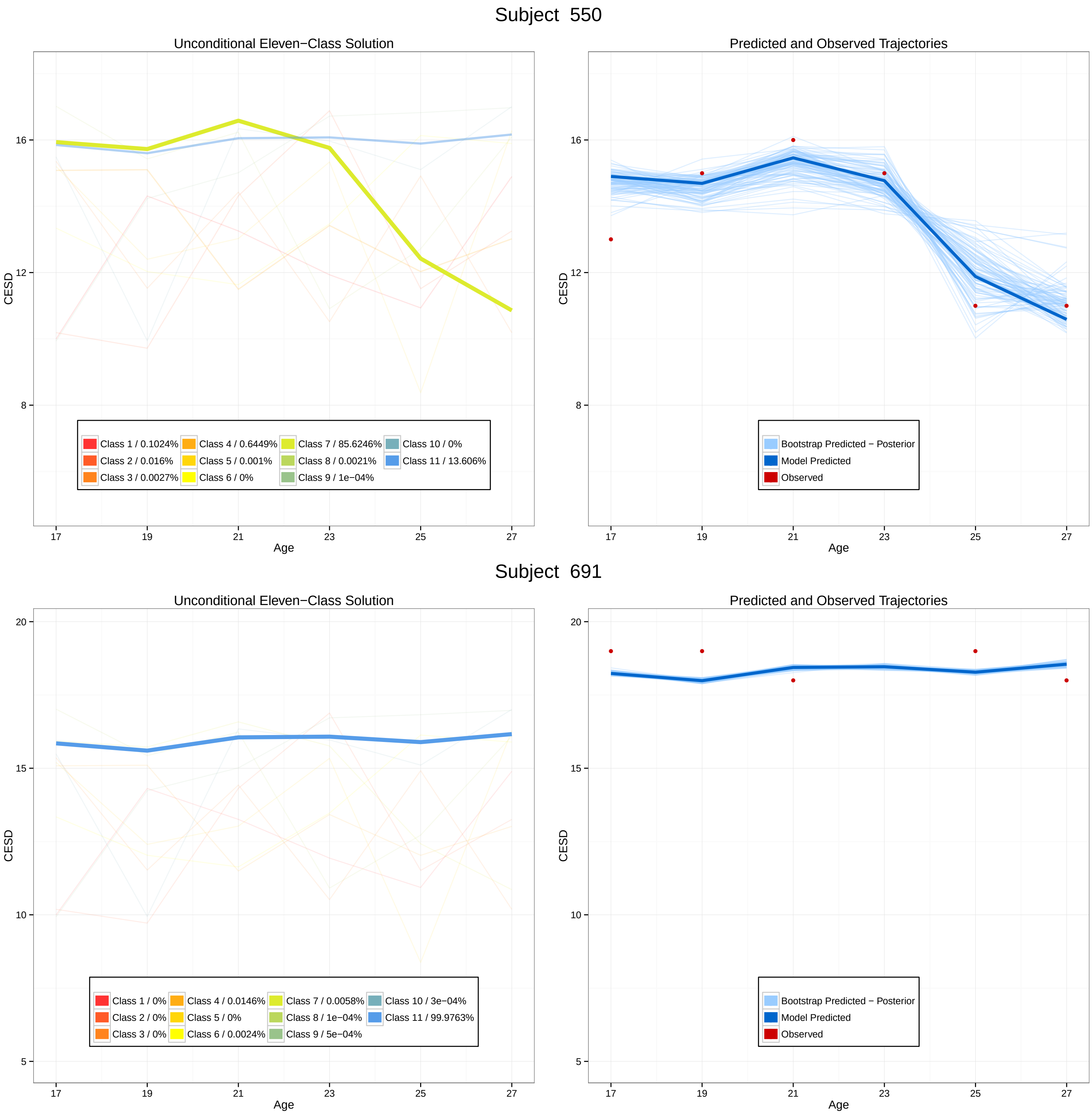


Figure 6: Predicted trajectories given gender and college attendance

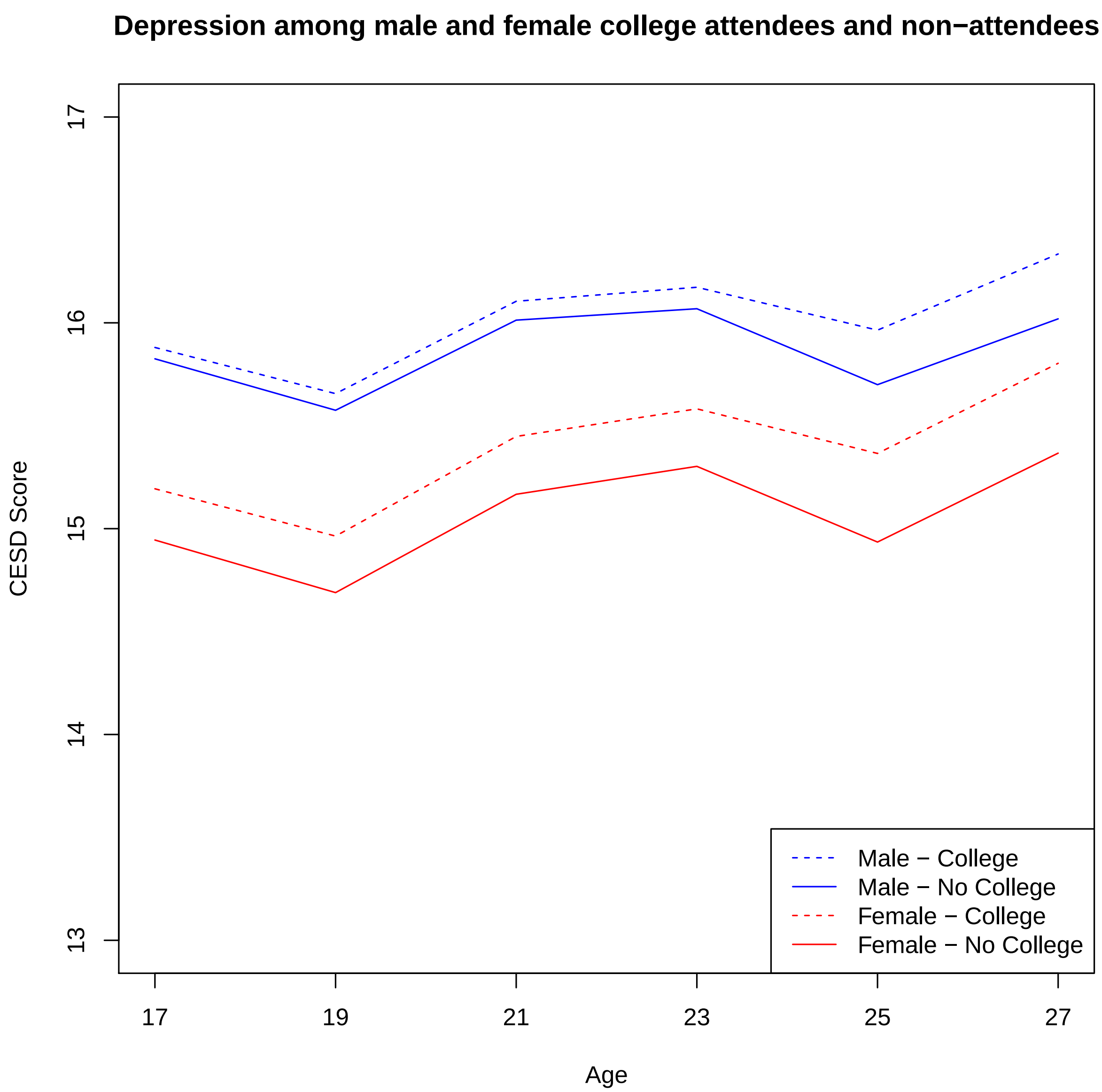
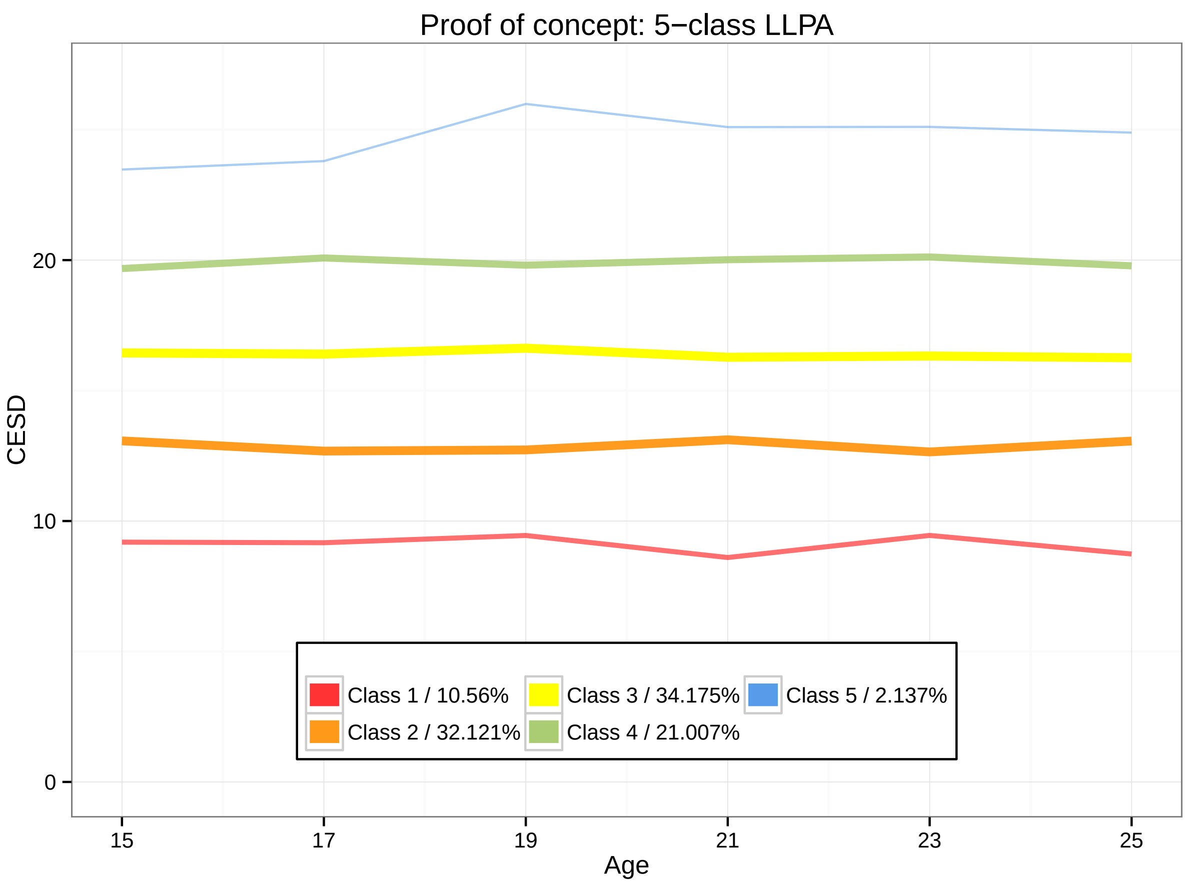


Figure 7: Predicted trajectories for proof-of-concept simulation



REFERENCES

Bartholomew, D., & Knott, M. (1999). *Latent class models and factor analysis.* London: Arnold.

Bartzokis, G., Cummings, J. L., Sultzer, D., Henderson, V. W., Nuechterlein, K. H., & Mintz, J. (2003). White matter structural integrity in healthy aging adults and patients with Alzheimer disease: a magnetic resonance imaging study. *Archives of Neurology, 60*(3), 393–398.

Bauer, D. J. (2005). A semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(4), 513–535.

Bauer, D. J. (2007). 2004 Cattell Award Address: Observations on the use of growth mixture models in psychological research, *42*(4), 757–786.

Bauer, D. J., & Curran, P. J. (2003a). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*(3), 338–63.

Bauer, D. J., & Curran, P. J. (2003b). Overextraction of latent trajectory classes: Much ado about nothing? Reply to Rindskopf (2003), Muthén (2003), Cudeck and Henly (2003). *Psychological Methods, 8*(3), 384–393.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: potential problems and promising opportunities. *Psychological Methods, 9*(1), 3–29.

Bauer, D. J., & Shanahan, M. J. (2007). Modeling complex interactions: Person-centered and variable-centered approaches. In T. D. Little, J. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 255–284). Mahwah, NJ: Lawrence Erlbaum Associates.

Bergman, L., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, *9*(2), 291–319.

Blanco, C., Okuda, M., Wright, C., Hasin, D. S., Grant, B. F., Liu, S.-M., & Olfson, M. (2008). Mental health of college students and their non-college-attending peers: results from the national epidemiologic study on alcohol and related conditions. *Archives of General Psychiatry, 65*(12), 14–29.

Blozis, S. A., & Cudeck, R. (1999). Conditionally linear mixed-effects models with latent variable covariates. *Journal of Educational and Behavioral Statistics, 24*(3), 245–270.

Browne, M. W. (1993). Structured latent curve models. *Multivariate analysis: Future directions*, *2*, 171–197.

Browne, M. W., & Du Toit, S. H. (1991). Models for learning data. *Best methods for the analysis of change*, 47–68.

Casswell, S., Pledger, M., & Pratap, S. (2002, November). Trajectories of drinking from 18 to 26 years: identification and prediction. *Addiction*(11), 1427–1437.

Castro-Schilo, L., & Ferrer, E. (2013). Comparison of nomothetic versus idiographic-oriented methods for making predictions about distal outcomes from time series data. *Multivariate Behavioral Research*, *48*(2), 175–207.

Chassin, L., Pitts, S., & Prost, J. (2002, February). Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: Predictors and substance abuse outcomes. *Journal of Clinical and Consulting Psychology*, *70*(1), 67–78. doi: 10.1037// 0022-006X.70.1.67

Coffman, D. L., & Millsap, R. E. (2009). Structural equation modeling : A evaluating latent growth curve models using individual fit statistics evaluating latent growth curve models using individual fit statistics. *Structural Equation Modeling: A Multidisciplinary Journal*(April 2013), 37–41.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2006). Testing and probing interactions in hierarchical linear growth models. In C. Bergeman & S. Boker (Eds.), *The notre dame series on quantitative methodology* (Vol. 1, pp. 99–129). Mahwah, NJ: Lawrence Erlbaum Associates.

Degnan, K. A., Hane, A. A., Henderson, H. A., Moas, O. L., Reeb-Sutherland, B. C., & Fox, N. A. (2011, May). Longitudinal stability of temperamental exuberance and social-emotional outcomes in early childhood. *Developmental Psychology*, *47*(3), 765–780.

Dekker, M. C., Ferdinand, R. F., van Lang, N. D. J., Bongers, I. L., van der Ende, J., & Verhulst, F. C. (2007). Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. *Journal of Child Psychology and Psychiatry*, *48*(7), 657–666.

Dumenci, L., Achenbach, T. M., & Windle, M. (2011). Measuring context-specific and cross-contextual components of hierarchical constructs. *Journal of psychopathology and behavioral assessment*, *33*(1), 3–10.

Everson, S. A., Maty, S. C., Lynch, J. W., & Kaplan, G. A. (2002). Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. *Journal of psychosomatic research*, *53*(4), 891–895.

Gfroerer, J. C., Greenblatt, J. C., & Wright, D. A. (1997). Substance use in the US college-age population: differences according to educational status and living arrangement. *American Journal of Public Health*, *87*(1), 62–65.

Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, *24*(3), 229–252.

Grimm, K. J. (2007). Multivariate longitudinal methods for studying developmental relationships between depression and academic achievement. *International Journal of Behavioral Development*, *31*(4), 328–339.

Grimm, K. J., & Ram, N. (2009). A second-order growth mixture model for developmental research. *Research in Human Development*, *6*(2), 121–143. doi: 10.1080/15427600902911221

Grimm, K. J., Ram, N., & Estabrook, R. (2010). Nonlinear structured growth mixture models in mplus and OpenMx. *Multivariate Behavioral Research*, *45*(6), 887–909. doi: 10.1080/00273171.2010.531230

Grimm, K. J., Steele, J. S., Ram, N., & Nesselroade, J. R. (2013). Exploratory latent growth models in the structural equation modeling framework. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 568–591.

Hankin, B. L., Mermelstein, R., & Roesch, L. (2007). Sex differences in adolescent depression: Stress exposure and reactivity models. *Child development*, *78*(1), 279–295.

Holliday, E. G., McLean, D. E., Nyholt, D. R., & Mowry, B. J. (2009). Susceptibility locus on chromosome 1q23-25 for a schizophrenia subtype resembling deficit schizophrenia identified by latent class analysis. *Archives of General Psychiatry*, *66*(10), 1058.

Karp, I., O'Loughlin, J., Paradis, G., Hanley, J., & Difranza, J. (2005). Smoking trajectories of adolescent novice smokers in a longitudinal study of tobacco use. *Annals of epidemiology*, *15*(6), 445–52.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*(6), 593.

King, R., Gaines, L., Lambert, E., Summerfelt, W., & Bickman, L. (2000). The co-occurrence of psychiatric and substance use diagnoses in adolescents in different service systems: Frequency, recognition, cost, and outcomes. *Journal of Behavioral Health Services and Research*, *27*(4), 417–430.

Laird, N., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*(4), 963–974.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.

Lin, N., & Ensel, W. M. (1984). Depression-mobility and its social etiology: The role of life events and social support. *Journal of Health and Social Behavior*, 176–188.

Lorant, V., Delige, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic inequalities in depression: a meta-analysis. *American Journal of Epidemiology*, *157*(2), 98–112.

Lubke, G., & Muthen, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(1), 26–47.

Maggs, J. L., Ph, D., Schulenberg, J. E., & Ph, D. (2005). Trajectories of alcohol use during the transition to adulthood. , *28*(4), 195–201.

Marsh, H. W., Luedtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(2), 191–225.

McArdle, J. J. (2004). Latent growth curve analysis using structural equation modeling techniques. In S. M. B. e. al (Ed.), *The handbook of research methods in developmental psychology* (pp. 340–466). New York: Blackwell Publishers.

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

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York, NY: Wiley.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, *60*(1), 117–174.

Meredith, W., & Tisak, J. (1984, July). ”Tuckerizing” curves. *Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA*.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107–122.

Mitchell, J. E., Crosby, R. D., Wonderlich, S. A., Hill, L., le Grange, D., Powers, P., & Eddy, K. (2007, November). Latent profile analysis of a cohort of patients with eating disorders not otherwise specified. *International Journal of Eating Disorders*, *40*(S), S95–S98.

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218.

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112–117.

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables. In *New methods for the analysis of change* (pp. 291–313).

Muthén, B. (2006). The potential of growth mixture modelling. *Infant and Child Development*, *15*(6), 623–625.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcoholism, clinical and experimental research*, *24*(6), 882–91.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*(2), 463–9.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods*, *4*(2), 139–157.

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

Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. *Criminology*, *31*(3), 327–362.

Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: Fact or a useful statistical fiction? *Criminology*, *43*(4), 873–904.

Nesselroade, J. R. (1988). Some implications of the trait-state distinction for the study of development over the life span: The case of personality.

Nolen-Hoeksema, S. (1990). *Sex differences in depression*.

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society*, *185*, 71–110.

Pek, J., Losardo, D., & Bauer, D. J. (2011). Confidence intervals for a semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(4), 537–553.

Pek, J., & MacCallum, R. C. (2010). Case diagnostics in structural equation models: Illustrations and issues. *Multivariate Behavioral Research*, *45*(6), 1030–1030.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437–448.

Rabe-Hesketh, S., & Skrondal, A. (2008). Classical latent variable models for medical research. *Statistical Methods in Medical Research*, *17*, 5–32.

Radloff, L. S. (1977). The center for epidemiologic studies depression scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: family social environments and the mental and physical health of offspring. *Psychological Bulletin*, *128*(2), 330.

Schulenberg, J., Merline, A., Johnston, L., O'Malley, P., Bachman, J., & Laetz, V. (2005). Trajectories of marijuana use during the transition to adulthood: The big picture based on national panel data. *Journal of Drug Issues*, *35*(2), 255–279.

Schulenberg, J., OMalley, P., Bachman, J., Wadsworth, K., & Johnston, L. (1996, May). Getting drunk and growing up: Trajectories of frequent binge drinking during the transition to young adulthood. *Journal of Studies on Alcohol*, *57*(3), 289–304.

Schulenberg, J., Wadsworth, K., OMalley, P., Bachman, J., & Johnston, L. (1996). Adolescent risk factors for binge drinking during the transition to young adulthood: Variable- and pattern-centered approaches to change. *Developmental Psychology*, *32*(4), 659–674. (5th Bi-Annual Meeting of the Society-for-Research-on-Adolescence, San Diego, CA, Feb 10-13, 1994)

Schwarz, G. (1987). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Sher, K. J., Jackson, K. M., & Steinley, D. (2011). Alcohol use trajectories and the ubiquitous cat's cradle: cause for concern? *Journal of Abnormal Psychology*, *120*(2), 322.

Skrondal, A., & Rabe-hesketh, S. (2003). Some applications of generalized linear latent and mixed models in epidemiology : Repeated measures , measurement error and multilevel modeling. *Norsk Epidemiologi*, *13*(2), 265–278.

Skrondal, A., & Rabe-Hesketh, S. (2009). Prediction in multilevel generalized linear models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(3), 659–687.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic* *and advanced multilevel modeling, second edition* (Vol. 1). London: Sage.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience*, *24*(38), 8223–8231.

Sterba, S. K., & Pek, J. (2012). Individual influence on model selection. *Psychological Methods*, *17*(4), 582.

Titterington, D. M., Smith, A. F., Makov, U. E., et al. (1985). *Statistical analysis of finite mixture distributions* (Vol. 7). Wiley New York.

Tucker, J., Orlando, M., & Ellickson, P. (2003). Patterns and correlates of binge drinking trajectories from early adolescence to young adulthood. *Health Psychology*, *22*(1), 79–87.

Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, *23*(1), 19–23.

Vaughn, M. G., Perron, B. E., & Howard, M. O. (2007, December). Variations in social contexts and their effect on adolescent inhalant use: A latent profile investigation. *Drug and Alcohol Dependence*, *91*(2), 129–133.

Wade, T. D., Crosby, R. D., & Martin, N. G. (2006, December). Use of latent profile analysis to identify eating disorder phenotypes in an adult australian twin cohort. *Archives of General Psychiatry*, *63*(12), 1377–1384. (11th Annual Meeting of the Eating-Disorders-Research-Society, Toronto, Canada, Sep 29-Oct 01, 2005)

Wechsler, H., Dowdall, G., Davenport, A., & Castillo, S. (1995, July). Correlates of college-student binge drinking. *American Journal of Public Health*, *85*(7), 921–926.

Weitzman, E., Nelson, T., & Wechsler, H. (2003). Taking up binge drinking in college: The influences of person, social group, and environment. *Journal of Adolescent Health*, *32*(1), 26–35.

White, H., Labouvie, E., & Papadaratsakis, V. (2005). Changes in substance use during the transition to adulthood a comparison of college students and their noncollege age peers. *Journal of Drug Issues*, *35*(2), 281–305.

White, H., Padina, R., & Chen, P. (2002). Developmental trajectories of cigarette use from early adolescence into young adulthood. *Drug and Alcohol Dependence*, *65*(2), 167–178.

Windle, M., & Wiesner, M. (2004). Trajectories of marijuana use from adolescence to young adulthood: Predictors and outcomes. *Development and Psychopathology*, *16*(4), 1007–1027. (Michigan Symposium on Development and Psychopathology, Ann Arbor, MI, JUN 14-15, 2002)

Wolfe, J. H. (1970). Pattern clustering by multivariate mixture analysis. *Multivariate Behavioral Research*, *5*, 329–350.