Nonlinear Change Models in Heterogeneous Populations When Class Membership is Unknown: The Latent Classification Differential Change Model

Ken Kelley
Inquiry Methodology Program
Indiana University

***Abstract*— When unobserved heterogeneity exists in populations where the phenomenon of interest is governed by a functional form of change linear in its parameters, the *growth mixture model* (GMM) is extremely useful for modeling change conditional on latent class (Muthén, 2001a; Muthén, 2001b; Muthén, 2002). However, when the functional form of interest is nonlinear in its parameters, the GMM is not very useful because it is based on a system of equations linear in its parameters. The *latent classification differential change* (LCDC) model is proposed and developed so that unobserved heterogeneity can be modeled in populations where the phenomenon of interest is governed by functional forms of change nonlinear in its parameters. Due to the integration of nonlinear multilevel models and finite mixture models, neither or which have closed form solutions, analytic solutions do not generally exist for the LCDC model. Five methods of parameter estimation are developed and evaluated with a comprehensive Monte Carlo simulation study. The simulation showed that the parameters of the LCDC model can be accurately estimated with each of the proposed methods, and that the method of choice depends on the precise question of interest.**

***Index Terms*— longitudinal data analysis, analysis of change, growth modeling, growth mixture modeling, heterogeneous population, heterogeneous change, nonlinear growth models, nonlinear change models, functional form of growth, functional form of change**

I. Introduction

An overarching theme permeating the whole of the scientific enterprise is the study of change. Although true for science in general, understanding the process of change in developmental and learning contexts are especially important, as key component of each is change. The difficulties that arise when attempting to understand the process of change in general has motivated an abundant amount of methodological work. Even though so much time and effort has been dedicated to gaining a better understanding of the process of change, state-of-the-art methods are not adequate when unknown heterogeneity exists in a population where change is governed by some nonlinear functional form. When the process of change for some phenomenon of interest is governed by a nonlinear functional form, most development and learning processes being a special case of such a phenomenon, modern methods of analyzing change do not address the fact that latent classes of individuals might exist and that those individuals might have a different set of parameter values e.g., Browne & du Toit, 1991; van Geert, 1991; Cudeck, 1996.

When studying change, especially in the behavioral sciences where a multitude of interindividual variability exists, it is often reasonable to hypothesize that not all individual entities change in the same fashion. It may be the case that there are underlying classes (subpopulations) of entities, where the different classes each have a unique set of parameters governing the process of change or even follow different functional forms of change altogether. If such classes exist, standard methods of longitudinal data analysis applied to phenomenon following a nonlinear functional form may not yield meaningful results. The reflexive application of change models that have a single set of parameters that supposedly represents the population of individual entities may in fact represent few or none of the individual entities. When a single population is composed of two or more unknown subpopulations (i.e., latent classes), treating the individuals as representative of a single population, as in a standard change model, yields averaged results (the results are analogous to a weighted mean across latent classes).

In such a scenario the conclusions drawn are potentially misleading because each of the subpopulation distributions is contaminated with the other subpopulation(s). Basing parameter estimates on contaminated distributions is problematic because the estimates are not representative of the properties of either distribution. What would be ideal is if the parameter estimates of change could be conditional on the latent class. In doing so the resultant parameter estimates would be representative of the properties within each of the latent classes. Such a scenario is ideal because the parameter estimates are meaningful for the individual entities contained in each of the latent classes. Thus, questions of interest regarding change would best be applied within each of the classes (i.e., conditional on class membership) and not across all of the classes simultaneously.

In a series of works Muthén has proposed a general latent variable technique that encompasses the growth mixture model (GMM; Muthén, 2001a, 2001b, 2002; Muthén & Shedden, 1999). Within the GMM context individuals are

Kelley, K. (2006). Nonlinear change models in heterogeneous populations when class membership is unknown: The latent classification differential change model. *Proceedings of the International Conference on Development and Learning*, 5, Bloomington, IN.

nested within unknown and thus latent classes, and within these latent classes there is generally variability around the class specific population coefficients of change. The GMM model is analogous to a multiple group analysis, but group membership is itself unknown and must be estimated. Although Muthén's GMM is capable of combining variable-centered and pattern-centered approaches to studying change (Muthén & Muthén, 2000), the GMM is limited in several respects. The major limitations of the GMM is that it is built on a linear system of equations and thus it is necessarily limited when change models nonlinear in their parameters are of interest.

Recent work reflecting increased appreciation of models of change nonlinear in their parameters is beginning to gain ground in the behavioral sciences. Methodologists have been giving models nonlinear in the parameters increased attention because of the unrivaled benefits they provide compared with models linear in their parameters. Nonlinear models of change are models where parameters enter the mathematical function defining the trajectories in a nonlinear fashion. As has been delineated numerous times (e.g. Thissen & Bock, 1990; Browne & du Toit, 1991; Davidian & Giltinan, 1995; Cudeck, 1996; Pinheiro & Bates, 2000), models nonlinear in their parameters provide many benefits, especially for modeling behavioral phenomenon where there is generally some limited capacity (i.e., the rate of learning slows as the material is mastered). Nonlinear models have many advantages over their counterparts whose parameters enter linearly. For example, nonlinear models need not exhibit the unlimited growth or decay characteristic of linear models, oftentimes nonlinear models of change require fewer parameters than do linear models for change that is asymptotic and/or sigmoidal ("S-shaped"), and the parameters of nonlinear models oftentimes have "real-world" interpretations. Furthermore, nonlinear models many times conform to a researchers theory about the phenomenon of interest as it changes over time (e.g., when knowledge has been mastered the trend asymptotes to some maximal value; the learning curve may be shallow at first, become steep, and then become shallow again; the point at which curvature changes from concave up to concave downward [i.e., an inflection point] may be a theoretically meaningful and predicted value, etc.).

As has been historically understood in the behavioral sciences since Allport (1937; but see Lamiell, 1998, for theoretical clarification of Windelband's, 1921, originating work on the distinction between idiographic and nomothetic conceptualizations), an idiographic approach to understanding behavior involves focusing on the individual and acknowledges that individuals may be qualitatively or quantitatively different from one another, whereas the nomothetic approach to understanding behavior attempts to establish universal laws that generalize over individuals (see also Dunn, 1994). Idiographic and nomothetic conceptualizations of behavior are traditionally considered antithetical (Dunn, 1994, p. 377). Because idiographic and nomothetic conceptualizations are often thought of as being mutually exclusive, potential benefits of combining the approaches in an integrated fashion often go unrecognized. What has not often been realized in practice, however, is that combining the idiographic and nomothetic conceptualizations of behavior into a single model of change may lead to more realistic and powerful models of change than either approach alone.

The proposed work attempts to fully integrate idiographic nomothetic approaches in an analysis of change context for models of change nonlinear in their parameters. In so doing, variable-centered and pattern-centered techniques are combined in an effective and unified way, where the focus is not on either a completely homogeneous or a completely heterogeneous population, but rather on some combination of the two. If one were to accept the premise that some behavior or pattern of behavior is best explained by relying on both idiographic and nomothetic conceptualizations, the method of data analysis should be flexible enough to allow such a rich theory to be incorporated into the statistical model. The latent classification differential change (LCDC) model is developed for such purposes when individuals follow nonlinear trajectories and multiple subgroups of individuals exist (this model subsumes many commonly used change models as special cases).

The LCDC model is conceptually similar to Muthén's GMM, however, the model flexibility and estimation procedures that define the LCDC model for classification and parameter estimation are fundamentally different from the GMM. There is greater flexibility in the functional forms of change that can be modeled with the LCDC model, because unlike Muthén's GMM the LCDC model is not restricted to functional forms linear in their parameters. When a process changes in a nonlinear fashion over time, and unknown heterogeneity exists, the LCDC model provides a conceptually appealing method of parameter estimation that is more flexible than current methods for understanding unknown heterogeneity in the context of longitudinal data analysis.

The potential problems of using nomothetic models that assume a homogeneous population on idiographic phenomenon where latent classes exist are tremendous. Obtaining parameter estimates that are from a mixed distribution generally implies that neither of the mixtures (latent classes in this case) are accurately represented by the parameter estimates. Basing important decisions on the results of a study that used a nomothetic model for data that consists of latent classes can be very problematic. Such decisions are unknowingly being based on biased estimates. The potential bias can be negligible or it can be tremendous, determining on the extent of the mixture and the separation of the latent classes.

With the LCDC model, many potentially interesting and important questions can be addressed. A few of those that fit into this framework are given below:

1) What functional form of change describes academic change (e.g., linear, quadratic, asymptotic regression, logistic, Gompertz, etc.) for some entity of interest (e.g., regions of the country, states, districts, etc.)?
2) What are the parameter estimates of the hypothesized functional form of change (e.g., at what age does an asymptotic value seem to be reached, what is the rate of linear change)?
3) How much variability exists for the parameter estimates of the hypothesized functional form of change (e.g., does the slope parameter different for different states or are states changing at the same rate)?
4) Do 'clusters' of entities seem to exist, where the variability within cluster is much less than the variability between 'clusters' (e.g., do unknown latent classes of states exist where their change is relatively homogeneous)?
5) Do quantitative differences in change parameters exist for the same functional form (especially for functional forms nonlinear in their parameters), leading to clusters of entities being well represented by a set of fixed parameter values (with potentially unique effects around the parameters) that differ in value from other clusters of entities?
6) Do qualitative differences in change exist, where the functional form of change is different for different clusters (i.e., latent classes) of entities (e.g., some states follow a straight-line change model while others follow an asymptotic regression change model)?

The six research questions listed above are very rich and, if answered, would provide a great deal of information to researchers whose questions are not readily solvable given the current state of research methodology on the analysis of change. Because the answers to these questions would provide so much information about change, yet the fact that they do not have unambiguous solutions provides a major shortcoming in the literature. Questions 4, 5, and 6 from above are especially well suited to researchers working within the areas of development and learning, because variability across individuals will not always be fully explainable by random variation around a single set of population parameters alone (i.e., there more be more than a single set of fixed effects with unique effects around these parameter estimates). There may be unknown but real differences among the individuals in terms of their change parameters because of latent classes (Question 4). If such heterogeneity exists, it would be beneficial to explore why such differences exist. Furthermore, individuals might be falling behind the typical trajectory and may benefit from an intervention or some sort of remedial attention of a particular learning task. Thus, clusters of individuals might change in a different fashion because of quantitative differences (Question 5) in change or because of qualitative differences (Question 6) in change.

Quantitative differences in change would occur when a single functional form of change exists to describe state change over time, but different latent classes exist, each with their own unique set of parameter values. Qualitative differences in change would exist when the functional form describing change is somehow different across the latent classes in the population, and thus the functionals are not a nested versions of one another. For example, the highest achieving individuals might seem to be governed by an asymptotic change curve (i.e., a negative exponential), whereas those failing to achieve adequate change might be governed by a logistic change curve. The importance of getting the model correct and appropriately carrying out the model is important if the change parameters are to be meaningful. In situations where unknown heterogeneity exists yet is ignored, parameter estimates will be based on mixed distributions, which leads to bias in estimated parameter values.

II. The Latent Classification Differential Change Model

Because of the deficiency of statistical methods currently available for modeling longitudinal data when interest lies in latent classes that each have their own potentially unique set of parameters from a change model nonlinear in its parameters, a new model for longitudinal data analysis is proposed. The model is a *latent classification* model because unknown classes of individuals are assumed to exist within a heterogeneous population. It is a *differential change* model because it is assumed that the different classes of individuals are governed by a unique set of unknown parameters and thus change differentially over time. The model explicitly acknowledges the fact that different classes of individuals potentially have their own unique set of parameter values, with potential constraints across class, and allows within class variability around the class specific fixed effects. Conceptually the LCDC model is an extension of the GMM. The LCDC model is a conceptual extension of the GMM because the LCDC model allows nonlinear change models to be carried out in heterogeneous populations when class membership is unknown.

The LCDC model is predicated on a heterogeneous population (like the GMM), where the parameters of change differ across latent classes and are governed by a change model nonlinear in its parameters (unlike the GMM). These interindividual differences in change are in part a function of class membership and in part a function of individual uniqueness. In this sense, the LCDC fulfills the ideals of integrating the idiographic and nomothetic conceptualizations of behavior into a single unified model. The core LCDC model is nonlinear in its parameters, where a special case is a model linear in its parameters. An assumption of the LCDC model is that a finite number of classes exist, each with a unique set of parameter values. Although the general model allows a unique class specific value for each model parameter, parameter invariance can be imposed for selected parameters across some or all of the classes (with some of the proposed methods of estimation). Conceptually, the LCDC is a combination of the FMM and the nonlinear MLM. In particular, the theory of the LCDC combines both methodologies so that a finite number of classes can each be represented by a nonlinear MLM with unique class specific parameters.

The LCDC model assumes that individuals within a class have trajectories of change that are relatively homogeneous, whereas trajectories of change across class are relatively heterogeneous. Of course, because the LCDC model assumes that interindividual differences in change are a function of both class membership and individual uniqueness, the degree of within group homogeneity is relative to the degree of across group heterogeneity. This is the case because unique effects are assumed to exist around the fixed effects within each class, while different parameter values are allowed across class. Following the formal definition of the LCDC model, five estimation procedures are proposed as ways to obtain parameters estimates of the LCDC model. The accuracy of the methods is then explored via a comprehensive Monte Carlo simulation study.

A. Defining the Latent Classification Differential Change Model

Let y_{it} be the outcome variable of the ith individual ($i = 1, \ldots, N$) at the tth measurement occasion ($t = 1, \ldots, T_i$), a_{it} be the value of some nonstochastic time dependent basis at the tth measurement occasion for the ith individual, and x_{im} be the mth time invariant predictor variable ($m = 1, \ldots, M$) for the ith individual. Further let $\mathbf{y}_i = [y_{i1}, \ldots, y_{iT_i}]'$ be a vector of length T_i of the observed scores for the ith individual, $\mathbf{a}_i = [a_{i1}, \ldots, a_{iT_i}]'$ be the vector of length T_i of the time dependent basis for the ith individual, $\mathbf{x}_i = [x_{i1}, \ldots, x_{iM}]'$ be the vector of length M of time invariant predictor variables for the ith individual, and $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]'$ be the N by M matrix of predictor variables for all individuals. The functional form governing the trajectory of change for all individuals is denoted $f(\mathbf{a})$, where $\mathbf{a}$ is a generic representation of the time dependent basis with the function $f(\cdot)$ having P parameters. Given $f(\mathbf{a})$, let $\boldsymbol{\theta}_i = \left[\theta_{i1}, \ldots, \theta_{ip}\right]'$ be a P length vector of true change coefficients defining the trajectory of interest for the ith individual and $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N]'$ be the N by P matrix of individual change coefficients. The vector of length N that identifies which of the G classes from a mutually exclusive "crisp set" the ith individual is a member is denoted $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_N]'$. A crisp class in this context is where individuals are assigned to one and only one class and thus no partial membership exists.

Given $f(\mathbf{a})$, $\boldsymbol{\Theta}$, and $\boldsymbol{\gamma}$, let $f(\mathbf{a}_g)$ be a special case of $f(\mathbf{a})$ for the gth class with $\pi_g \left(\sum_{g=1}^{G} \pi_g = 1\right)$ being the proportion of the population who are members of the gth class and $n_g \left(\sum_{g=1}^{G} n_g = N\right)$ being the number of individuals in the sample who are members of the gth class. The functional form of $f(\mathbf{a}_g)$ is thus equal to or a special case of $f(\mathbf{a})$, however some if not all of the P parameters of change in $f(\mathbf{a}_g)$ are specific to the gth class. Conditional on $\boldsymbol{\theta}_i$, γ_i, and $\boldsymbol{\mu}_g$, where $\boldsymbol{\mu}_g = [\mu_{1g}, \ldots, \mu_{Pg}]'$ is the gth class specific population mean vector for the P fixed effect parameters of change, $\mathrm{E}[\boldsymbol{\upsilon}_i] = \mathbf{0}$ with variance $\boldsymbol{\Sigma}_g$ and has a P dimensional multivariate normal distribution, where $\boldsymbol{\upsilon}_i = [\upsilon_{i1}, \ldots, \upsilon_{iP}]'$ are the individual specific unique effects defined as $\upsilon_{ip} = \theta_{ip} - (\mu_{gp}|\gamma_i)$.

The LCDC model implies that the coefficients of change for the individuals conform to the following probability density function:

$$\mathrm{d}(\boldsymbol{\theta}_i, \mathbf{x}_i; \boldsymbol{\Psi}) = \sum_{g=1}^{G} \pi_g \phi_P(\boldsymbol{\theta}_i, \mathbf{x}_i; \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g), \qquad (1)$$

where $\boldsymbol{\theta}_i$ and $\mathbf{x}_i$ are taken together, ϕ_P represents a P dimensional multivariate normal probability density function, and $\boldsymbol{\Psi} = \left[\boldsymbol{\Psi}'_1, \ldots, \boldsymbol{\Psi}'_G\right]'$, where $\boldsymbol{\Psi}_g = \left[\pi_g, \boldsymbol{\mu}'_g, \mathrm{vech}(\boldsymbol{\Sigma}_g)'\right]'$ that may or may not be saturated with vech$(\cdot)$ being an operator that stacks the columns of a symmetric matrix by leaving out elements above the main diagonal. Equation 1, which defines the density of the heterogeneous population of change coefficients, is analogous to the general density of a mixture distribution (e.g., McLachlan & Basford, 1988; McLachlan & Peel, 2001)

What is of interest in the LCDC framework is not literally the probability density function of $\boldsymbol{\theta}_i$ given $\mathbf{x}_i$, rather what is of interest are the parameters from a nonlinear MLM with multiple classes of individuals. However, class membership is unknown and must itself be estimated. The desire is to have a multiple group nonlinear MLM, where rather than having a known grouping structure the grouping structure is latent and must be estimated, implies a MLM of the form:

$$\mathbf{y}_{ig} = f(\mathbf{A}_{ig}\boldsymbol{\beta}_g + \mathbf{Z}_{ig}\boldsymbol{\upsilon}_{ig}, \mathbf{x}_{ig}) + \boldsymbol{\varepsilon}_{ig}, \qquad (2)$$

where g is unknown and must be estimated. Thus, the LCDC model is conditional on the unknown class memberships (i.e., g), which themselves must be estimated.[1] The class identification vector, $\boldsymbol{\gamma}$, can be estimated with the optimal rule of classification with a FMM whose probability density is defined by Equation 1. The optimal rule for classification simply states that an individual entity should be classified into the class that the entity is most likely to belong (McLachlan & Basford, 1988, pp. 11, 45–46; sometimes this is called the Bayes rule, Anderson, 1984, chapter 6). Note the one-to-one relationship between $\boldsymbol{\theta}_i$, the ith individual's change

[1]Note that if g in Equation 2 where know it could be replaced by j and the LCDC model would reduce to the nonlinear MLM.

coefficients, in Equation 1 and $\boldsymbol{\upsilon}_{ig}$, the unique effects associated with the ith individual whom is in the gth class from Equation 2.

Each of the classes in the general LCDC model of Equation 2 potentially has a unique set of fixed effects, $\boldsymbol{\beta}_g$, a unique covariance matrix of errors, $\boldsymbol{\Sigma}_{\varepsilon_g}$, and has a unique covariance matrix for the unique effects, $\boldsymbol{\Sigma}_{\upsilon_g}$. Without imposing constraints on the model parameters, either by fixing parameters to specified values or imposing parameter invariance, there are potentially a large number of parameters that can be estimated. Momentarily supposing each of the individuals is measured at the same T measurement occasions, there are potentially GP fixed effect parameters, $G(T(T+1))2^{-1}$ error covariance parameters, and $G(P(P+1))2^{-1}$ unique covariance parameters. Thus, with no parameter constraints, there are a total of $G[p+(T(T+1)+p(p+1))2^{-1}]$ parameters defining the LCDC model. Of course, in order for the model not to be under-identified, parameter constraints must be imposed. Thus, not all possible parameters can be free to vary, as doing so yields an inestimable model where there are more equations than unknowns (see Loehlin, 1998, for a discussion of these issues in a different, yet related, context). A reasonable constraint reduces a large number of parameter if $\sigma^2_{\varepsilon g}\mathbf{I}$ is the error structure for each of the classes, and even more constrained with $\sigma^2_{\varepsilon 1} = \sigma^2_{\varepsilon 2} = \ldots = \sigma^2_{\varepsilon G}$. The potential flexibility of the model necessarily implies more parameters than either a standard MLM or a standard FMM. Given the difficulties that often exist for estimation of complicated FMMs and especially for nonlinear MLMs, it is no surprise that estimation of the LCDC model can be difficult. The next section discusses how likelihood estimation theoretically applies to the parameters of the LCDC model.

B. Parameters of the LCDC Model

Because the LCDC model combines ideas from the FMM with ideas from the general nonlinear MLM, the likelihood function for the LCDC model can be written as

$$\mathcal{L}(\mathbf{Y},\mathbf{X};\boldsymbol{\Omega}) = \prod_{i=1}^{N} \mathrm{d}(\mathbf{y}_i,\mathbf{x}_i;\boldsymbol{\Omega}) \tag{3}$$

where $\mathcal{L}$ represent the likelihood function and $\boldsymbol{\Omega}$ is the R length vector of parameter values defining the particular LCDC model. The parameter vector $\boldsymbol{\Omega}$ can be written in a general form:

$$\begin{gathered}\boldsymbol{\Omega} = [\pi_1,\ldots,\pi_{G-1},\boldsymbol{\beta}'_1,\ldots,\boldsymbol{\beta}'_G,\\ \mathrm{vech}(\boldsymbol{\Sigma}_{\varepsilon_1})',\ldots,\mathrm{vech}(\boldsymbol{\Sigma}_{\varepsilon_G})',\mathrm{vech}(\boldsymbol{\Sigma}_{\upsilon_1})',\ldots,\mathrm{vech}(\boldsymbol{\Sigma}_{\upsilon_G})']',\end{gathered} \tag{4}$$

where a unique $\boldsymbol{\beta}_g$ is the vector of fixed effects for the gth class, $\boldsymbol{\Sigma}_{\varepsilon_g}$ is the covariance matrix of errors for the gth class, and $\boldsymbol{\Sigma}_{\upsilon_g}$ is the covariance matrix for unique effects for the gth class.

An obvious question that arises is how the likelihood function of Equation 3 relates to the probability density function defined in Equation 1. The vector of group specific population means contained in $\boldsymbol{\mu}_g$ from Equation 1 is analogous to the fixed effect change parameters contained in $\boldsymbol{\beta}'_g$ from Equation 4. Furthermore, the class specific population covariances contained in $\boldsymbol{\Sigma}_g$ from Equation 1 are analogous to the class specific covariances contained in $\boldsymbol{\Sigma}_{\upsilon_g}$ from Equation 4. Knowing this, Equation 1 can be rewritten so that it is consistent with the notation used for the likelihood function of the LCDC model defined in Equation 3:

$$\mathrm{d}(\boldsymbol{\theta}_i,\mathbf{x}_i;\boldsymbol{\Psi}) = \sum_{g=1}^{G} \pi_g \phi_P(\boldsymbol{\theta}_i,\mathbf{x}_i;\boldsymbol{\beta}_g,\boldsymbol{\Sigma}_{\upsilon_g}). \tag{5}$$

III. Methods of Parameter Estimation for the LCDC Model

In large samples when all assumptions are met and the model is correctly specified, parameter estimation by full information maximum likelihood provides many benefits and is generally the optimal method from a statistical point of view. However, for many complicated models, especially those with nonlinear unique effects (Pinheiro & Bates, 2000, section 7.2.1), the likelihood function may be difficult or impossible to write analytically given the current state of knowledge and the limitations of statistical theory. If the likelihood function cannot be analytically derived, an approximate likelihood function can sometimes be used. Rather than maximizing the likelihood function, a reasonable way to proceed is to make use of an approximate likelihood function. As Davidian and Giltinan describe in the context of nonlinear MLMs, "the analytical intractability of likelihood inference has motivated many approaches based on approximations" for nonlinear MLMs (2003, p. 403). Thus, given the current limitations of likelihood estimation and inference, carrying out nonlinear MLMs is almost always based on approximate procedures. As a result, parameter estimates based on approximate likelihood functions are themselves necessarily approximate.

Given the approximate nature of parameter estimates from nonlinear MLMs, the estimation of the parameters of the LCDC model is necessarily constrained. Evidence of this is understandable given that a special case of the LCDC model is a standard nonlinear MLM when $G = 1$. In such a situation, the "mixture" is actually a single homogeneous population and the LCDC model is reduced to a MLM. Because the present work does not attempt to improve or propose estimation procedures for the nonlinear MLM, but rather relies on estimation procedures commonly used for carrying out the model, a known shortcoming of the LCDC model (like nonlinear MLMs in general) is that estimation of the parameters is based on estimation methods that do not have the optimal properties of full information maximum likelihood.

Five different estimation procedures are proposed that are conceptually reasonable and theoretically statistically sound methods for estimating the parameters of the LCDC model.

A large scale simulation study was conducted to evaluate the effectiveness of the proposed procedures. A very general summary is that the methods were shown to work well in realistic situations, with the most appropriate estimation method depending on the precise question of interest the researcher is interested in evaluating. The methods of estimation and their effectiveness will be discussed.

IV. Conclusion

When heterogeneity exists in an analysis of change context, ignoring the heterogeneity introduces bias into the obtained results. In such a situation the parameter estimates that are supposed to represent a single homogeneous population instead represent a mixture of homogeneous subpopulations. This potentially translates into a model of change that is meant to represent everyone with a single trajectory but in actuality may fit no one. When theory or data suggest the existence of latent classes in an analysis of change context, combining a MLM or LGC framework with the FMM is necessary to ensure the parameter estimates are conditional on the latent classes. In cases where the functional form governing change is linear in its parameters, the GMM provides a powerful and elegant model to help understand change. However, in cases where the functional form of change is governed by a nonlinear functional form, which is probably more often than not true in the behavioral sciences, use of the LCDC model of change is appropriate. The LCDC model of change and its estimation methods can address a diverse set of questions arising from data having a wide range of complicated structures. The hope is that by explicitly modeling unknown heterogeneity in populations where the functional form of change has parameters that enter nonlinearly will help better address the questions asked by applied researchers who are interested in understanding, and not simply analyzing, change.

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Henry Hold and Company.

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Browne, M. W., & du Toit, S. H. C. (1991). Models for learning data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions.* Washington, DC: American Psychological Association.

Cudeck, R. (1996). Mixed-effects models in the study of individual differences with repeated measures data. *Multivariate Behavioral Research*, *31*(3), 371–403.

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. New York, NY: Chapman & Hall.

Davidian, M., & Giltinan, D. M. (2003). Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, *8*(4), 387–419.

Dunn, J. G. H. (1994). Toward the combined use of nomothetic and idiographic methodologies in sport psychology: An empirical example. *The Sport Psychologist*, *8*, 376–392.

Lamiell, J. T. (1998). 'Nomothetic' and 'idiographic:' Contrasting Windelband's understanding with contemporary usage. *Theory & Psychology*, *8*(1), 23–38.

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering*. New York, NY: Marcel Dekker, Inc.

McLachlan, G. J., & Peel, D. (2001). *Finite mixture models*. New York, NY: John Wiley & Sons, Inc.

Muthén, B. (2001a). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Muthén, B. (2001b). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L. M. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (p. 291-322). Washington, DC: American Psychological Association.

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variables-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*(6), 882–891.

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

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-Plus*. New York, NY: Springer.

Thissen, D., & Bock, R. D. (1990). Linear and nonlinear curve fitting. In A. von Eye (Ed.), *Statistical methods in longitudinal research, Volume ii: Time series and categorical longitudinal data* (pp. 289–318). Orlando, FL: Academic Press.

van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*(1), 3–52.

Windelband, W. (1921). *An introduction to philosophy*. London: Unwin.