National Symposium on Family Issues

Susan M. McHale
Paul Amato
Alan Booth
Editors

Emerging Methods in Family Research

Springer

National Symposium on Family Issues

Series Editors
Alan Booth and Susan M. McHale

For further volumes:
http://www.springer.com/series/8381

Susan M. McHale • Paul Amato • Alan Booth
Editors

Emerging Methods in Family Research

Springer

Editors
Susan M. McHale
Department of Human Development
& Family Studies
The Pennsylvania State University
University Park
Pennsylvania
USA

Paul Amato
Department of Sociology
The Pennsylvania State University
University Park
Pennsylvania
USA

Alan Booth
Department of Sociology
The Pennsylvania State University
University Park
Pennsylvania
USA

ISSN 2192-9157 ISSN 2192-9165 (electronic)
ISBN 978-3-319-01561-3 ISBN 978-3-319-01562-0 (eBook)
DOI 10.1007/978-3-319-01562-0
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013949340

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Across a range of disciplines, including psychology, sociology, education, political science and medicine, research and theory target families as central to the well-being of their members and to the well-being of the communities and larger societies in which families are embedded. The significance of families for the health of individuals and communities demands scholars' best efforts to illuminate how family roles, relationships and dynamics operate and how they influence family members.

This volume is predicated on the idea that advances in research on families will rely on innovations in design, measurement, data collection and data analysis that allow researchers to capture the multi-level complexities of family systems. Methods for studying families are often drawn from research focused on individuals. A theme throughout this volume is whether and to what extent the same kinds of methods can be applied across levels of analysis—from the individual, to the dyad, to larger family groups. In chapters throughout this volume, authors consider whether and how methods from research focused on individuals can be applied, can be modified, and are challenged when family relationships and family influences are the focus of study.

The contributions to *Emerging Methods in Family Research* are based on papers presented at the 20th Annual Penn State Symposium on Family Issues held in October, 2012. This edited volume is the culmination of two days of stimulating presentations and discussions organized around four topics: (1) strategies for quantitative analysis of variation and change in families, (2) approaches to analyzing families as systems, (3) measuring "the family" and family dynamics, and (4) new directions in the implementation and evaluation of family-focused social policies and preventive interventions.

Overview of this Volume

This volume is organized by these four topical areas. In contrast to other volumes in the decades-long Family Symposium series, our focus on methods meant that many of these chapters were written by researchers who do not self- identify as family scholars, but rather, are known for their methodological expertise. These methodologists accepted the editors' invitation to apply their work and ideas to the study of families. The four sections of this volume each include two or three chapters that address the topical area in distinct ways, often from different disciplinary perspectives. The last chapter in each section is an integrative discussion by a family scholar who was charged with distilling the range of ideas, information, and techniques described in the session's papers toward providing insights on how novel methods could be used to advance the work of family scholars. The volume concludes with an integrative chapter by two young scholars.

Chapters in Part I focus on best methods for capturing variation and change in family processes and influences on individual family members. Family structure and processes are dynamic, responsive both to changes in individual family member's development, as well as to pressures emanating from outside the family, which are also continually in flux. Chapters by Jay Teachman, Professor of Sociology at Western Washington University, by Nilam Ram, Associate Professor of Human Development at Penn State, and colleagues, and by Si-Miin Chow, Associate Professor of Human Development at Penn State, and colleagues, focus on different timescales for studying variation and change, timescales that reflect different kinds of research questions and require different kinds of analytic methods. In the final chapter in this section, Andrew Fuligni, Professor of Psychiatry and Behavioral Sciences at UCLA, outlines some of the contributions to our understanding of family processes and family influences that can come from sophisticated analyses of variation and change, and he considers how the benefits of collecting "repeated data" balance against the costs. For researchers interested in why one might use the models introduced in the three opening chapters, Fuligni's application of each model to the case of family sleep patterns conveys the distinctive insights that can emerge from each approach.

Family scholars have long embraced the metaphor of families as systems, yet empirical research targeting systems dynamics remains very rare. In Part II, chapters by Robin Gauthier and James Moody, both sociologists at Duke University, and by Mark Cummings and co-authors Kathleen Bergman and Kelly Kuznicki, psychologists at Notre Dame University, focus on methods for characterizing family systems and capturing their dynamics. In his integrative discussion, Robert Emery, Professor of Community Psychology at the University of Virginia, reinforces and elaborates on the important conceptual and theoretical work that must be accomplished if family researchers are to make full use of a systems approach, and he offers new ideas toward this end.

At a general level, measurement is "concerned with what can be observed, the conditions under which observations are made, and how observations are recorded for future analysis and consideration" (Amato, Chap. 11, p. 179). In Part III, chapters by Carolyn Tucker Halpern from the Department of Maternal & Child Health along with Kathleen Mullan Harris from the Department of Sociology and epidemiologist Eric Whitsel, all at University of North Carolina, Chapel Hill, by Joshua Smyth, Professor of Biobehavioral Health and co-author Kristin Herron from Penn State University, and by Thomas Weisner, Professor of Psychiatry and Anthropology at UCLA, describe distinct approaches to measuring family dynamics and their correlates. As with other dimensions of methods considered in this volume, most family research relies on measurement approaches that were developed to study individuals, and these chapters include consideration of approaches for and challenges to moving from the individual to the dyad and group levels of analysis in measuring family processes and influences. In the concluding chapter, Paul Amato, Professor of Sociology at Penn State, considers some of the strengths of these approaches to measurement and provides examples of how each might be applied to address novel questions about family processes and influences. Amato also reminds us of the challenges of defining “the family” in determining strategies for its measurement.

Recent national efforts have been directed at promoting the translation of science to application and practice as well as improving the quality of programs and policy through a focus on evaluation. Although a stronger emphasis on applying research in evidence-based programs and policies is welcome, the development, implementation and evaluation of programs and policies for families face unique challenges. The chapters in Part IV by Carol Metzler, from the Oregon Research Institute and colleagues, by Quinn Moore and Robert Wood from Mathematica Policy Research, and by Linda Collins, Professor of Human Development at Penn State, highlight new approaches to optimal design, implementation and evaluation of the effects of family programs and policies and consider some of the challenges that need to be overcome toward these ends. In the final, integrative chapter of this section, Greg Duncan, from the School of Education at the University of California Irvine, identifies a number of “best practices” in family-focused evaluation and policy research.

As is the tradition in the Family Symposium series, the final chapter of the volume was written by two scholars in the early stages of their careers as family researchers, Melissa Lippold, from Human Development and Family Studies, and Catherine McNamee, from the Population Research Institute at Penn State. Their charge was to bring to bear their distinct disciplinary backgrounds—in human development and demography, respectively—on the ideas and insights conveyed during the four sessions of the conference. Lippold and McNamee identify five themes that cut across chapters in this volume: approaches to defining “family” and capturing its complexities, assessing change and variation in families, the challenges inherent is studying families, the importance of keeping in sight the “big picture,” and the significance and special considerations involved in family research that is aimed at improving public health. Lippold and McNamee conclude with their thoughts about opportunities and challenges facing the next generation of family scholars.

Acknowledgments

The editors are grateful to the many organizations at Penn State that sponsored the 2012 Symposium on Family Issues and this resulting volume, including the Population Research Institute; the Children, Youth and Families Consortium; the Prevention Research Center; the Methodology Center; the departments of Sociology, Human Development and Family Studies, Psychology, Biobehavioral Health, Anthropology, and Labor Studies; and the Women's Studies Program. The editors also gratefully acknowledge essential core financial support in the form of a five-year grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), R13- HD048150, as well as advice from Regina Bures and Rosalind Berkowitz King of NICHD. The ongoing support of all of these partners has enabled us to attract excellent scholars from a range of backgrounds and disciplines—-the sort of group on whom the quality of the series depends.

A lively, interdisciplinary team of scholars from across the Penn State community meets with us annually to generate symposia topics and plans and is available throughout the year for brainstorming and problem solving. We appreciate their enthusiasm, intellectual support, and creative ideas. In the course of selecting speakers, symposium organizers consult with a wide range of people at other universities, at NICHD, and at other organizations in order to identify highly qualified scholars to participate in the symposium. We also sincerely thank Diane Felmlee, Jennifer Graham, and Wayne Osgood for presiding over symposium sessions.

The efforts of many individuals went into planning the 2012 symposium and producing this volume. We are especially grateful for the assistance of the administrative staff in the Population Research Institute and Social Science Research Institute at Penn State, including Sherry Yocum, Donna Panasiti, Angela Jordan, and Miranda Bair. Finally, neither Symposium nor this volume would have been possible without Carolyn Scott, whose organizational skills, commitment, and attention to the many details that go into developing an engaging conference and an edited book series make it possible for us to focus on the ideas.

Contents

Contributors

Paul Amato, Ph.D. Department of Sociology and Crime, Law, & Justice, The Pennsylvania State University, University Park, PA 16802, USA

Kathleen N. Bergman, M.A. Department of Psychology, Notre Dame University, Notre Dame, IN, USA

Sy-Miin Chow, Ph.D. Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA

Linda M. Collins, Ph.D. The Methodology Center, Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA

David Conroy, Ph.D. Department of Kinesiology, The Pennsylvania State University, University Park, PA, USA

E. Mark Cummings, Ph.D. Department of Psychology, Notre Dame University, Notre Dame, IN, USA

Shawna Doerksen, Ph.D. Department of Recreation, Park and Tourism Management, The Pennsylvania State University, University Park, PA, USA

Greg J. Duncan, Ph.D. School of Education, University of California, Irvine, CA, USA

Robert E. Emery, Ph.D. Department of Psychology, Center for Children, Families & the Law, University of Virginia, Charlottesville, VA, USA

Andrew J. Fuligni, Ph.D. Department of Psychiatry & Biobehavioral Science, University of California, Los Angeles, CA, USA

Robin Gauthier Department of Sociology, Duke University, Durham, NC, USA

Carolyn Tucker Halpern, Ph.D. Department of Maternal & Child Health, University of North Carolina Chapel Hill, Chapel Hill, NC, USA

Kathleen Mullan Harris, Ph.D. Department of Sociology, University of North Carolina Chapel Hill, Chapel Hill, NC, USA

Kristin E. Heron, Ph.D. Survey Research Center, The Pennsylvania State University, University Park, PA, USA

Kelly A. Kuznicki Department of Psychology, Notre Dame University, Notre Dame, IN, USA

Melissa A. Lippold, Ph.D. Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA

Erika S. Lunkenheimer, Ph.D. Department of Human Development & Family Studies, Colorado State University, Fort Collins, CO, USA

Whitney I. Mattson, M.S. Department of Psychology, University of Miami, Miami, FL, USA

Catherine B. McNamee, Ph.D. Population Research Institute, The Pennsylvania State University,University Park, PA, USA

Daniel S. Messinger, Ph.D. Department of Psychology, University of Miami, Miami, FL, USA

Carol W. Metzler, Ph.D. Oregon Research Institute, Eugene, OR, USA

James Moody, Ph.D. Department of Sociology, Duke University, Durham, NC, USA

Quinn Moore, Ph.D. Mathematica Policy Research, Princeton, NJ, USA

Nilam Ram, Ph.D. Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA

Julie C. Rusby, Ph.D. Oregon Research Institute, Eugene, OR, USA

Matthew R. Sanders, Ph.D. Parenting and Family Support Centre, University of Queensland, Brisbane, QLD, Australia

Mariya Shiyko, Ph.D. Department of Applied Counseling and Educational Psychology, Northeastern University, Boston, MA, USA

Joshua M. Smyth Departments of Biobehavioral Health & Medicine, The Pennsylvania State University, University Park, PA, USA

Jay Teachman, Ph.D. Department of Sociology, Western Washington University, Bellingham, WA, USA

Thomas S. Weisner, Ph.D. Departments of Psychiatry and Anthropology, Los Angeles, CA, USA

Eric A. Whitsel, M.D., M.P.H. Departments of Epidemiology and Medicine in the Gillings School of Global Public Health and School of Medicine, University of North Carolina, Chapel Hill, NC, USA

Robert G. Wood, Ph.D. Mathematica Policy Research, Princeton, NJ, USA

Part I
Family Development and Change

Chapter 1
Latent Growth Curve Models with Random and Fixed Effects

Jay Teachman

The availability of longitudinal data on families is increasingly common. Data sets such as the Panel Study of Income Dynamics, the various National Longitudinal Studies, AddHealth, and Fragile Families allow researchers numerous opportunities to observe and model family-related processes and outcomes as they evolve over time. Accordingly, a number of statistical procedures have been developed to model repeated observations of families and individuals. Two common alternatives are random-effects models (REM) and fixed-effects models (FEM) (Allison 2005; Bollen and Brand 2010) Within the general random-effects framework, latent growth curve models (LGCM) are a useful extension because they allow researchers to explicitly model the *trajectory* of change in an outcome (Lyons and Sayer 2005) In this chapter, using a structural equation modeling (SEM) approach, I demonstrate that LGCMs can also be estimated within a fixed-effects framework. In addition, I show that time-constant covariates, which are generally modeled on the inter-subject level in LGCMs, can be modeled on the intra-subject level. These models are illustrated using data on marital status, education, and body mass index (BMI) for 1761 men taken from four waves (1992, 1996, 2000, 2004) of the 1979 National Longitudinal Survey of Youth (NLSY). Umberson et al. (2009). provide a discussion of the relationship between BMI, marital status, and other covariates. Finally, I show that LGCMs can be used to model paired data using information on marital satisfaction gathered from 218 continuously-married couples in the Early Years of Marriage Project (EYMP).

Traditional Random- and Fixed-Effects Models for Longitudinal Data

In much of the family literature the most common procedure for examining repeated observations on individuals or other units of observation is a REM or a FEM. Many statistical packages easily allow estimation of REMs and FEMs. STATA (XTREG)

J. Teachman (✉)
Department of Sociology, Western Washington University, Bellingham, WA, USA
e-mail: Jay.Teachman@wwu.edu

S. M. McHale et al. (eds.), *Emerging Methods in Family Research,*
National Symposium on Family Issues 4, DOI 10.1007/978-3-319-01562-0_1,

and SAS (PROC GLM for fixed effects and PROC MIXED for random effects) are popular options. It is also possible to estimate REMs and FEMs using structural equation models (Bollen and Brand 2010; Teachman et al. 2001) using programs such as Mplus, EQS, AMOS, or PROC CALIS in SAS. SEMs allow researchers to explicitly model or manipulate co-variances, and as I demonstrate, SEMs allow hybrid models mixing both fixed and random effects, as well as extensions to models such as LGCMs. Accordingly, in this chapter all models are presented and estimated as SEMs.

To begin the discussion and fix ideas, consider the following REM:

$$y_{it} = \alpha_t + \beta_{\mathbf{yxt}}\mathbf{X_{it}} + \beta_{\mathbf{yzt}}\mathbf{Z_i} + \beta_{y\eta t}\eta_i + \varepsilon_{it} \tag{1.1}$$

y_{it} is the value of the dependent variable for the i^{th} case at time t; α_t is an intercept term at time t; $\mathbf{X_{it}}$ is a vector of time-varying covariates for the i^{th} case at time t; $\beta_{\mathbf{yxt}}$ is a vector of coefficients indicating the effects of $\mathbf{X_{it}}$ on y_{it}; $\mathbf{Z_i}$ is a vector of time-constant covariates for the i^{th} case; $\beta_{\mathbf{yzt}}$ is a vector of coefficients indicating the effects of $\mathbf{Z_i}$ on y_{it}; η_i is a scalar indicating all of the latent time-constant factors affecting y_{it}; $\beta_{y\eta t}$ is the coefficients linking the latent factor η_i to y_{it} at time t (here all values of this vector are set equal to 1.0); and ε_{it} is a random disturbance for the i^{th} case at time t with $E(\varepsilon_{it}) = 0$ and $E({\varepsilon^2}_{it}) = {\sigma^2}_{\varepsilon t}$. It is assumed that ε_{it} is uncorrelated with $\mathbf{X_{it}}$, $\mathbf{Z_i}$, and η_i, that $COV(\varepsilon_{it},\varepsilon_{it}) = 0$ for $t \neq s$, and that η_i is uncorrelated with $\mathbf{X_{it}}$ and $\mathbf{Z_i}$. This is the default REM estimated by most software products in which the effects of the time-varying variables are constrained to be constant across time, as are the variances of the error terms. This is also the REM that many, if not most, researchers using longitudinal data report. The key assumption for the purposes of this chapter is that η_i is uncorrelated with the included covariates. Accepting this assumption means that the standard errors of the coefficients are adjusted for clustering but are not adjusted for unmeasured covariates that may be correlated with both the dependent variable and the covariates.

A FEM can be written as:

$$y_{it} = \alpha_t + \beta_{\mathbf{yx}}\mathbf{X_{it}} + \beta_{y\eta}\eta_i + \varepsilon_{it} \tag{1.2}$$

The $\mathbf{Z_i}$ are now dropped from the equation because they are assumed to be included in the time-constant latent variable η_i. The key assumption here is that the model allows η_i to be correlated with the $\mathbf{X_{it}}$. Note that even though one loses the ability to obtain an estimate of the impact of specific time-constant variables on the outcome, their effects are still controlled by including η_i in the model. In a FEM, the effects of the time-varying variables are washed of any effect linked to unmeasured time-constant factors. Without prior evidence, I suggest that assuming zero covariance between latent time-constant factors and the included covariates is risky. If a correlation exists, and it is not modeled, the estimated parameter estimates will be biased, either upward or downward. The FEM avoids this issue.

If the REM does not include $\mathbf{Z_i}$ then it is simply a restricted version of the FEM that constrains the covariances between η_i and $\mathbf{X_{it}}$ to be zero. In other words, the models are nested and can be compared via a standard likelihood ratio test. If the

Table 1.1 Results from estimating random- and fixed-effects models for BMI

	Random effects	Fixed effects
Cohabiting	0.264**	0.280**
Married	0.485**	0.475**
Highest grade completed	−0.112**	0.087
T1 intercept	27.847**	25.171**
T2 intercept	28.729**	26.032**
T3 intercept	29.566**	26.855**
T4 intercept	29.894**	27.175**
LR chi-square	682.13	651.58
Df	53	41
RMSEA	0.082	0.092
BIC	286.03	345.16

** $p < 0.05$

REM includes $\mathbf{Z_i}$ but the FEM does not, then the models are not nested. On the other hand, if the REM includes $\mathbf{Z_i}$ and the FEM also includes $\mathbf{Z_i}$ (by constraining the covariance between $\mathbf{Z_i}$ and η_i to equal zero) the two models are once again nested (the result is a hybrid REM/FEM). Bollen and Brand (2010) provide an overview of these points within the context of a general model for panel data.

Results from estimating these two models using the NLSY data on BMI are shown in Table 1.1. The NLSY data consist of 1761 observations for men in 1992, 1996, 2000, and 2004. Cases with missing data were deleted, as were men with BMI values greater than 50. Marital status is time-varying and is measured as married, cohabiting, and other. Highest grade completed is time varying and indicates the highest year of schooling completed by the respondent. A time-constant indicator of race/ethnicity measured as Black, Hispanic, and other is included. PROC CALIS in SAS was used to generate these estimates. The estimates are identical to models estimated using XTREG in STATA. According to the REM, when compared to men not in a union, both cohabitors and married men are heavier, and men with more education have lower values of BMI. The FEM shows similar estimates of the effects of marital status, but the effect of highest grade completed has changed signs and is no longer statistically significant, indicating that its effect can be attributed to the common latent factor. The intercept terms indicate that for both the REM and the FEM there is a tendency for BMI to increase over time. A variety of fit statistics are provided for each model: LR chi-square, Root Mean Square of Approximation (RMSEA), and Bayesian Information Criterion (BIC). Differences between the LR chi-square values suggest that the FEM is a better choice than the REM ($X = 30.55$, 12 df). The BIC value for the FEM is much larger than the BIC value for the REM however, indicating some ambiguity in whether the FEM should be favored. In large part, this is a common occurrence because FEMs use additional degrees of freedom (by allowing non-zero covariances between the latent term and the time-varying covariates), and BIC penalizes models that use more degrees of freedom. Overall, the fit statistics do not indicate well-fitting models though. In particular, BIC values should be negative for models that fit the data well.

Latent Growth Curve Models for Longitudinal Data

The lack of fit for either the REM or FEM in Table 1.1 suggests that another model specification is in order. I argue that in the case of a variable like BMI a LGCM is appropriate. LGCMs are appropriate when the outcome variable being considered follows a trajectory of change across time (i.e., does not randomly shift across time). A simple LGCM can be expressed as follows (ignoring time-constant variables):

$$y_{it} = \beta_{\mathbf{yxt}}\mathbf{X_{it}} + \beta_{y\eta 0t}\eta_{oi} + \beta_{y\eta 1t}\eta_{1i} + \varepsilon_{it} \tag{1.3}$$

η_{oi} is a latent factor indicating initial values of BMI with slopes, $\beta_{y\eta 0t}$, constrained to equal 1, and η_{1i} is a second latent factor with slopes, $\beta_{y\eta 1t}$, indicating change in BMI over time. All other terms and assumptions are defined as in Eq. 1.1 with the caveat that the latent factor is now represented by two terms. Most researchers call η_{oi} the intercept (beginning or initial value of the outcome) and η_{1i} the slope of the model (change across time from the initial value of the outcome). In a LGCM therefore, there are two latent components rather than one as is the case in a REM. Similar to a REM, one latent factor (intercept) describes a stable component across time. (Slopes are constrained to unity.) The second latent factor allows variation from this stable component over time and can be thought of as representing the rate of change across time. This is the factor that models structured change across time.

LGCMs (as well as REMs and FEMs) are hierarchical linear models (HLM) in that there are two levels of variation represented: within-subject and between-subject. The $\mathbf{X_{it}}$ represent within-subject variation, whereas η_{oi} and η_{1i} represent between-subject variation. Because they vary between subjects, both η_{oi} and η_{1i} can be represented as functions of other time-constant covariates. This point is demonstrated later in this chapter. Many applications of REMs and FEMs ignore the fact that variation in outcomes occurs both within and between respondents.

A graphic representation of this model is shown in Fig. 1.1. Note that as shown this is a random-effects LGCM because both latent terms are assumed to be independent of any covariates. Also note that the slopes for the second latent term representing change in BMI are fixed at 0, 1, 2, and 3 to reflect a linear trajectory of gains in BMI. (Alternative specifications are possible.) The model shown is known as a conditional LGCM because the latent terms are estimated conditional on the effects of marital status and education (and vice versa). If marital status and education were not included in the model and each BMI included an error term instead, the model would be an unconditional LGCM. For the purpose of establishing a baseline, I estimate an unconditional LGCM using the NLSY data on BMI. The resulting value for η_{oi} is 26.80 and the resulting value for η_{1i} is 0.693. These values can be thought of as the average value of the intercept and slope, respectively. That is, if all 1761 cases were plotted, the average starting value for BMI would be 26.80 and the average slope indicating gain in BMI over time would be 0.693. Of course, these average values also have standard errors because they are not fixed across individuals. In this case, the standard error for the intercept term is 0.099, and the standard error for the slope term is 0.023. In both cases, a simple t-test indicates that there is statistically

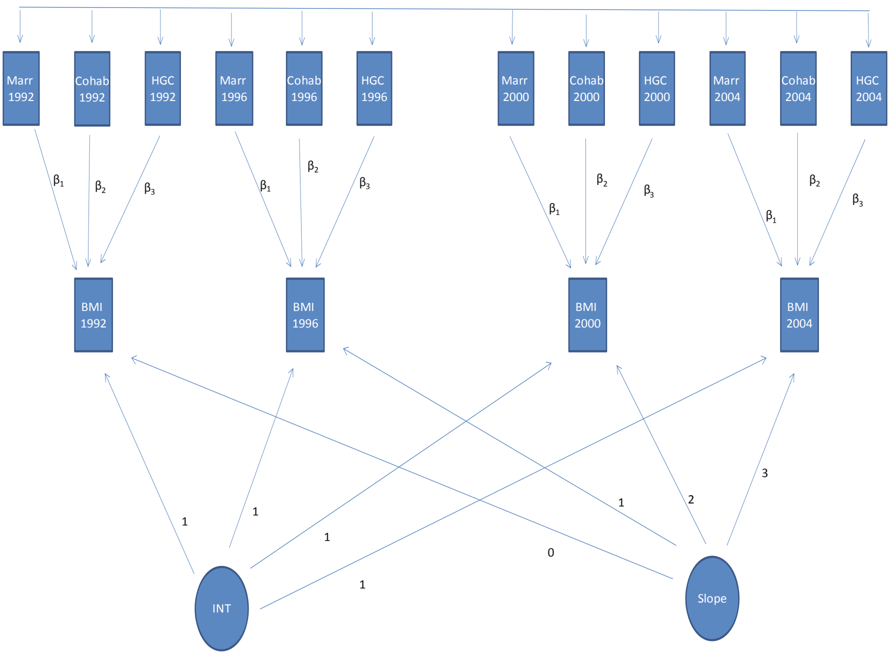

Fig. 1.1 Simple latent growth curve model with marital status and highest grade of schooling completed as time-varying covariates

significant variation in both initial value of BMI and rate of change across time (slope). The model fit statistics for the unconditional model are 182.64/8 for the LR Chi-square, 0.111 for RMSEA, and 122.85 for BIC.

The fit statistics for the unconditional LGCM are not particularly good and suggest that the model can be improved. Accordingly, I estimate the conditional LGCM represented in Fig. 1.1. The model fit statistics for this model are shown in Table 1.2 for Model A. These values indicate that this model is a better fit to the data. In particular, RMSEA is much lower (0.045 vs. 0.111), and BIC is now negative (−153.75). The parameter estimates for this model are shown in Table 1.3. The latent intercept has a value of 28.044 and the slope estimate is 0.697. Thus, the addition of the time-varying covariates does not dramatically alter estimates of these basic parameters. Yet, the fact that Model A fits the data better than an unconditional model indicates that net of the latent growth factors, marital status and education significantly affect BMI. Compared to the coefficients for traditional REM shown in Table 1.1, the effects of being married (0.515) and the effects of education (−0.118) are similar; however, the effect of cohabitation (0.135) is much smaller and not statistically significant. This latter result suggests that the effect of cohabitation as estimated in the traditional REM can be attributed to latent growth in BMI over time. That is, beyond the tendency for men to become heavier as they age, cohabitation does not appear to be related to BMI.

Table 1.2 Model fit statistics for various latent-growth curve models: NLSY data on male BMI

Model	LR chi-square/df	RMSEA	BIC
A. LGC REM	242.36/53	0.045	−153.75
B. LGC REM + quadratic	102.77/49	0.025	−263.44
C. LGC FEM	198.58/29	0.058	−18.16
D. LGC FEM + quadratic	61.98/25	0.020	−124.87
E. hybrid LGC + quadratic	83.63/45	0.022	−252.68
F. hybrid LGC + quadratic + mediated race/ethnicity	87.11/49	0.021	−279.10
G. hybrid LGC + quadratic + direct race/ethnicity equal slopes	101.02/51	0.024	−280.13
H. hybrid LGC + quadratic + direct race/ethnicity unequal slopes	72.18/45	0.019	−264.13
I. hybrid LGC + quadratic + direct race/ethnicity unequal slopes + unequal variances	38.59/42	0.000	−275.30
J. hybrid LGC + quadratic + direct race/ethnicity unequal slopes + unequal variances + level 1 unequal slopes	30.96/33	0.000	−215.67
K. hybrid LGC + quadratic + mediated race/ethnicity + unequal variances	53.81/46	0.010	−289.98

Some researchers might stop here after concluding that marriage and education affect BMI but that cohabitation does not. However, there are important extensions to Model A. One possible extension is to consider non-linear changes in BMI over time. There are many ways to allow for non-linear change in the outcome variable but perhaps the most parsimonious is to model a quadratic rate of change (by adding a quadratic latent term to Eq. 1.3). In this case, the additional latent construct allows for a quadratic change in the slope, in which each slope is just the square of the linear change in slope. (i.e., Each of the paths from the latent quadratic slope construct is the square of the corresponding latent linear slope construct). As shown in Table 1.2, this model (Model B) fits the data better than a model with only a linear term. RMSEA is now 0.025 (vs.0.045), and the value of BIC is more negative (−263.44 vs. −153.75). Table 1.3 shows the parameter estimates for this model. The intercept term is 28.171, the linear slope is 1.115, and the quadratic slope is −0.139, all statistically significant. The positive linear slope and negative quadratic slope indicate that BMI tends to increase over time but at a diminishing rate. The coefficients for marital status and education remain similar to those estimated for Model A, and the effect of cohabitation remains non-significant.

A Latent Growth Curve Fixed-Effects Model for Longitudinal Data

Another extension of the LGC REM model is to consider a LGC FEM model. As indicated in Fig. 1.1, in a LGC REM there are no covariances allowed between the latent terms and any of the time-varying covariates. If these covariances are allowed

Table 1.3 Estimated components of various latent-growth curve models: NLSY data on male BMI

Model components	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H	Model I	Model J	Model K
Manifest variable equations:											
Cohabiting	0.135	0.148	0.010	0.024	0.175	0.176**	0.164	0.187*	0.187*		0.176*
Time 1										0.367**	
Time 2										0.198	
Time 3										0.121	
Time 4										0.242	
Married	0.515**	0.490**	0.533**	0.483**	0.483**	0.499**	0.495**	0.498**	0.493**		0.494**
Time 1										0.602**	
Time 2										0.502**	
Time 3										0.434**	
Time 4										0.383**	
HGC	−0.118**	−0.137**	0.067	0.003	0.058	0.078	0.078	0.078	0.068		0.063
Time 1										0.061	
Time 2										0.081	
Time 3										0.082	
Time 4										0.050	
Black							1.010**				
Time 1								0.896	0.885**	0.883**	
Time 2								0.903**	0.900**	0.932**	
Time 3								1.373**	1.370**	1.339**	
Time 4								1.461**	1.451**	1.441**	
Hispanic							1.569**				
Time 1								1.563**	1.547**	1.530**	
Time2								1.450**	1.441**	1.480**	
Time 3								1.787**	1.779**	1.767**	
Time 4								1.678**	1.663**	1.659**	
Latent Variable Equations:											
Intercept	28.044**	28.171**	25.585**	26.376**	25.545**	24.838**	24.800**	24.836**	24.984**	24.994**	24.987**
Black						0.850**					0.857**

Table 1.3 (continued)

Model components	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H	Model I	Model J	Model K
Hispanic						1.519**					1.520**
Slope	0.697**	1.115**	0.682**	1.097**	1.091**	1.032**	1.088**	1.045**	1.035**	0.637*	1.009**
Black						0.215**					0.213**
Hispanic						0.068					0.068
Quadratic slope		−0.139**		−0.136**	−0.136**	−0.135**	−0.135**	−0.140**	−0.136**	−0.031	−0.128**
Error BMI1	2.024**	1.703**	2.021**	1.702**	1.703**	1.703**	1.702**	1.689**	0.890*	0.889	0.898*
Error BMI2	2.024**	1.703**	2.021**	1.702**	1.703**	1.703**	1.702**	1.689**	1.536**	1.535**	1.551**
Error BMI3	2.024**	1.703**	2.021**	1.702**	1.703**	1.703**	1.702**	1.689**	1.870**	1.868**	1.883**
Error BMI4	2.024**	1.703**	2.021**	1.702**	1.703**	1.703**	1.702**	1.689**	2.228**	2.239**	2.245**
Error intercept	15.493**	14.971**	15.843**	15.178**	15.364**	15.070**	15.073**	15.084**	15.782**	15.768	15.772**
Error slope	0.527**	1.671**	0.526**	1.660**	1.656**	1.652**	1.656**	1.687**	2.391**	2.377*	2.363**
Error quadratic slope		0.141**		0.141**	0.141**	0.141**	0.141**	0.145**	0.160**	0.157**	0.156**
Covariances:											
Intercept—slope	0.097	0.758**	−0.058	0.758**	0.763**	0.737**	0.733**	−0.015	−0.024	−0.029	−0.015
Intercept—quadratic slope		−0.297**		−0.295**	−0.300**	−0.299**	−0.300**	−0.128	−0.125	−0.122	−0.128
Slope—quadratic slope		−0.392**		−0.389**	−0.390**	−0.391**	−0.890**	−0.527**	−0.537**	−0.532**	−0.527**
Intercept—time 1 HGC					−1.582**	−1.470**	−1.469**	−1.470**	−1.422**	−1.388*	−1.421**
Intercept—time 2 HGC					−1.528**	−1.420**	−1.420**	−1.420**	−1.370**	−1.335**	−1.369**
Intercept—time 3 HGC					−1.551**	−1.449**	−1.448**	−1.449**	−1.398**	−1.363**	−1.398**
Intercept—time 4 HGC					−1.592**	−1.487**	−1.487**	−1.487	−1.438	−1.403	−1.438**

* $p < 0.10$
** $p < 0.05$

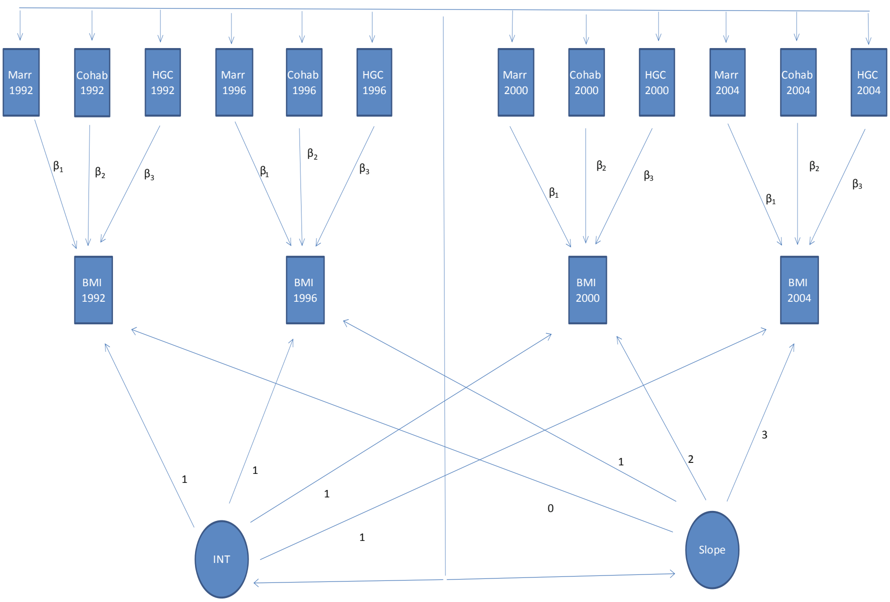

Fig. 1.2 Fixed-effects latent growth curve

to differ from zero then a LGC FEM model results. As far as I am aware, such a model has not been presented in the existing literature dealing with LGCMs even though the extension from an LGC REM is straightforward. A LGC FEM is illustrated in Fig. 1.2 where non-zero covariances between the latent intercept and slope terms and the time-varying covariates are allowed. The value in estimating a LGC FEM is that it will provide unbiased estimates of the effects of included covariates if they are correlated with unmeasured latent terms.

The fit statistics for the LGC FEM shown in Fig. 1.2 are presented in Table 1.2 (Model C). Compared to Model A, the RMSEA is larger (0.058 vs. 0.025) and BIC is less negative (−18.16 vs. −263.44). The parameter estimates for this model shown in Table 1.3 indicate an intercept term of 25.585 and a slope of 0.682, values similar to the LGC REM. The effects of cohabitation and marriage are also similar to those estimated in the LGC REM (although the non-significant effect of cohabitation is now larger), but the effect of education is no longer negative and statistically significant, indicating that its impact was due to covariation with the latent factor.

An improvement in model fit results when estimating a LGC FEM that allows a latent quadratic slope term. For the sake of parsimony, when estimating this model I allowed only the latent intercept and latent linear slope constructs to be correlated with the time-varying covariates. The fit statistics for Model D in Table 1.2 (LR chi-square 61.98/25, RMSEA = 0.020, and BIC = −124.87) indicate a better fit to the data than Model C. The parameter estimates shown in Table 1.3 indicate that, similar to Model C, the only time-varying covariate to affect BMI is being married. Although the LGC FEM in Model D with a quadratic term yields a better fit than the LGC FEM in Model C with only a linear term, the fit compared to the LGC REM with a quadratic term (Model B) is equivocal. Whereas the RMSEA is smaller (0.020 vs. 0.025), the difference in BIC statistics is considerable with the REM version possessing a much more negative BIC value (−263.44 vs. −124.87). The difference in BIC values suggests an elaboration of the LGC FEM that may lead to a better fitting model.

Specifically, the LGC FEM estimated in Model D includes a sizeable number of estimated covariances between the latent constructs and the time-varying covariates (even though the latent quadratic term was not allowed to covary with these covariates). As indicated earlier, BIC strongly penalizes models that estimate more parameters than necessary. If some of the covariances are statistically indistinguishable from zero, then valuable degrees of freedom are being wasted. When I examined the covariances estimated in Model D, I found that most of them were not distinguishable from zero. Indeed, the only covariances that were consistently significant involved schooling and the latent intercept term. Fit statistics for a LGC FEM allowing only these covariances between latent and observed terms are shown in Model E of Table 1.2. I term this a hybrid LGC because it involves only a subset of all possible covariances between the latent and observed variables. Compared to Model B, the LR chi-square value (83.63/45) and the RMSEA value (0.022) indicate a better fit to the data. The BIC value for this model (−252.68) is not quite as negative as the BIC value for the LGC REM with a quadratic term (−263.44) but is a significant improvement over earlier versions of the LGC FEM.

Parameter estimates for Model E are shown in Table 1.3. Also shown in Table 1.3 for Model E are the covariances between education and the latent intercept term. Each of the four covariances is statistically significant and negative. In other words, there are unmeasured factors that link having more education with lower body weight.

Model E suggests two important points with respect to the effects of the time-varying covariates. First, the effect of marital status is not substantially biased by failure to include covariances with the latent terms in the model. Thus, we can be more confident in our ability to state that being married is positively linked to BMI whereas cohabitation is not. Second, the effect of education is biased by the failure to include covariances with the latent terms in the model (here the latent intercept term).

Time-Constant Covariates in Latent Growth Curve Models with Fixed Effects

An extension often found in standard LGC REMs is to allow the latent constructs to be functions of time-constant observed variables. In other words, between respondent variation in the outcome under consideration can be modeled. Consider the following equations:

$$y_{it} = \beta_{\mathbf{yxt}} X_{it} + \beta_{\mathbf{y\eta 0t}} \eta_{oi} + \beta_{\mathbf{y\eta 1t}} \eta_{1i} + \varepsilon_{it} \tag{1.4}$$

$$\eta_{oi} = \alpha_{00} + \gamma_{\eta 0z} Z_i + \delta_{0i} \tag{1.5}$$

$$\eta_{1i} = \alpha_{10} + \gamma_{\eta 1z} Z_i + \delta_{1i} \tag{1.6}$$

α_{00} and α_{10} are constant terms; $\gamma_{\eta 0z}$ and $\gamma_{\eta 1z}$ are slopes; Z_i is a time-constant variable affecting the latent intercept and slope; δ_{0i} and δ_{0i} are error terms; and all other terms are as defined earlier. Equations 1.4–1.6 emphasize the hierarchical nature of the LGC framework, which uses information both within- and between-subjects. Within-subject variation is modeled as a function of the latent intercept and slope, as well as time-varying covariates. Between-subject variation in the latent intercepts and slopes is modeled as a function of variation on time-constant variables such as race or ethnicity.

Assuming that Z_i is a measure of race/ethnicity (Black and Hispanic), this model assumes that the effects of race/ethnicity on BMI are mediated by the latent intercept and slopes. Fit statistics for this model are presented in Table 1.2 (Model F). Compared to previous models, this model fits the data well with LR chi-square = 87.11/49, RMSE = 0 .021 and BIC = −279.10. Parameter estimates for this Model F are shown in Table 1.3. The results indicate that Blacks (0.850) and Hispanics (1.519) have higher initial levels of BMI, with the value for Hispanics being nearly twice as high as that for Blacks. The pace of increase in BMI is greater for Blacks (0.215) compared to Whites but does not differ for Hispanics. For the sake of simplicity, I did not allow race/ethnicity to affect the quadratic slope term.

Although not common in the literature, it is possible to allow time-constant variables to affect the time-varying dependent variable(s) directly. No special accommodations are necessary to do this in the LGC REM. In the LGC FEM, however, covariances between the latent terms and the time-constant variable must be set to zero in order for the model to be identified. The mediated model assumes that all of the effects of the time-constant variable are captured by its impact on the latent growth parameters. This assumption is strong and it may well be the case that the time-constant variable directly affects the outcome variables being examined and not the intercept and trajectory of change.

A model that includes both the direct and mediated effects of a time-constant variable is not identified. Thus, researchers will need to choose between the two. Fortunately, it has been shown that the mediated model is nested within the direct model (Stoel et al. 2004) This nesting means that a LR chi-square test can be used to determine whether the direct model is warranted. Table 1.2 presents fit statistics for a model allowing direct effects of race/ethnicity on BMI (Model G). This model constrains the effect of race/ethnicity to be constant across all four measurement points. The LR chi-square value (101.02/51) and the RMSEA (0.024) indicate that the mediated model should be preferred. (There is little difference between BIC values.) The parameter estimates for Model G in Table 1.3 indicate that the effect of being Black (1.010) or Hispanic (1.569) is to increase BMI at each point in time. Estimates of the latent growth parameters are similar to previous estimates.

As noted, Model G constrains the direct effects of race/ethnicity to be equal across all four time points. It may be plausible that these effects differ across time. Accordingly, Model H represents a model where the effects of race/ethnicity are free to vary over time. This model fits the data better than the mediated model according to the LR chi-square (72.18/45 vs. 87.11/49) and RMSEA (0.019 vs. 0.021). Because more degrees of freedom are being used though, the BIC value (−264.13 vs. −279.10) indicates a preference for the mediated model. Additional elaborations are possible. Model I in Table 1.2 extends Model H by allowing unequal variances for the error terms associated with each value of BMI. Model J continues relaxing assumptions in that the time-varying predictors are allowed to have effects on BMI that vary across time. Model fit statistics indicate that these models do not yield an improvement in fit to the data, and I do not discuss them further.

A final model (K) is shown in Table 1.2. In this model, the effects of race/ethnicity are mediated through the latent terms, and unequal error variances in BMI are allowed. The LR chi-square value (53.81/46) and RMSEA value (0.010) when compared to Model I do not indicate a better fit to the data. The value of BIC (− 289.98), however, indicates that this model may be the preferred option. The parameter estimates shown in Table 1.3 indicate parameter estimates very similar to those shown for Model F. Hispanics have the highest initial levels of BMI followed by Blacks and then Whites. Compared to Whites the rate of change in BMI is steeper for Blacks but not Hispanics. The freed variances for the four measures of BMI continue to indicate increasing random variation over time. In terms of marital status and education, being married or in a cohabiting relationship increases BMI, whereas there is no effect of schooling once a fixed-effects estimator is employed.

A Latent Growth Curve Model for Paired Data

I have outlined what I believe to be some valuable extensions of LGCMs for family data, using data on BMI and marital status taken from the National Longitudinal Study of Youth (NLSY). I now offer a further extension with a simple example. The extension I propose is for paired or matched family data. That is, data that refer to two or more members of the same family. Neale and McArdle (2000) have shown how LGCMs can be used to examine twin data. An empirical example of this procedure is provided by Hjelmborg et al. (2008) who demonstrate that genetic influences on rate of change in BMI are different from those affecting level of BMI. In essence this procedure takes advantage of the fact that multilevel models, including LGCMs, can be estimated simultaneously for multiple groups. A model is estimated simultaneously for several groups (e.g., identical twins, fraternal twins) and constraints are imposed on the various parameters of the model across groups in order to determine whether the models for particular twin groups differ from those of other twin groups.

Although this is a very useful approach and one that demonstrates the ability of the model to be estimated across groups, it does have limitations. In particular, such a model does not allow the parameters of one group to affect the parameters of another group, much like one would anticipate in a family group. Similarities are assumed to be a function of shared genetic potential rather than patterns of interaction. This limitation may make sense when examining twin data but is less useful in other circumstances. For example, consider the case of a married couple

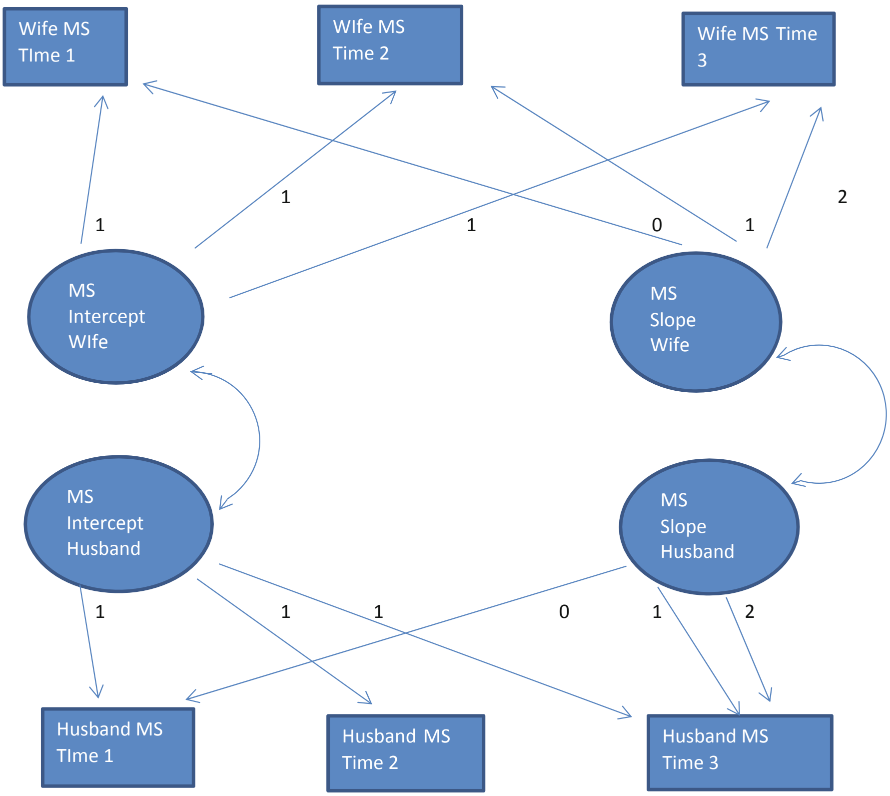

Fig. 1.3 Latent growth curve model for paired processes

and their respective paths of change in marital satisfaction over time. It may be the case that the trajectory of marital satisfaction in one spouse affects or is affected by the trajectory of the other spouse.

A generic example of this notion is illustrated in Fig. 1.3. Walker et al. (1996) provide an example of this sort of model using data on two variables (amount of caregiving performed and satisfaction with caregiving) measured for a single individual. In Fig. 1.3, I assume there is longitudinal information obtained on one variable (here marital satisfaction) for two related individuals (husband and wife). A very simple model is presented with no time varying covariates affecting the measured variable of interest. It is a simple extension of Fig. 1.3 to include time varying covariates and thus have the ability to estimate REM or FEM LGCMs in this framework as described earlier.

The most important components of the model in Fig. 1.3 for the purposes of matched processes are the relationships allowed between the intercepts and slopes across the two persons in the model. Two-way arrows are shown indicating no assumed directionality of effects (i.e., simply the covariance between the two terms).

With proper theoretical justification and appropriate specification, effects in each direction could be estimated. The arrow linking the intercepts indicates matching on the underlying value of the variable in question. Using the marital satisfaction of married couples, the covariance between intercepts can be interpreted to illustrate the degree of marital matching on relationship quality. The covariance between the slopes can be interpreted to illustrate the extent to which the paths of change in marital satisfaction within a married couple are linked.

I estimated such a model using four waves of data taken from the Early Years of Marriage Project (EYMP), (Veroff et al. 1986–1989) Data on marital satisfaction were collected in 1986, 1987, 1988, and 1989. Using information provided by 218 continuously-married couples, I computed a simple additive marital satisfaction scale based on four items: (1) how likely do you believe your marriage will be intact in five years, (2) how stable do you perceive your marriage to be, (3) have you ever considered leaving your marriage, and (4) how satisfied are you with your marriage. Higher scores indicate a greater degree of marital *dissatisfaction*.

The model estimated corresponds to that shown in Fig. 1.3 using four time points in the EYMP. I did not attempt to find a best fitting model, but I did compare two models that were otherwise identical. The first model constrained the covariances between the intercepts and slopes of husbands and wives to be equal to zero. The second model allowed these covariances to vary. The second model fitted the data much better ($X^2 = 120.35$ with 2 df). The covariance between the intercepts was 0.60, and the covariance between the slopes was 0.16. The slopes were positive (1.05 for wives and 0.87 for husbands) reflecting growing marital dissatisfaction over time. The positive correlation between the slopes for the spouses indicates that the pace at which one spouse's marital dissatisfaction grows influences the pace at which the other spouse's dissatisfaction grows. Failure to consider the joint influence of each spouse on the other would yield a biased estimate of the pace of change in marital dissatisfaction. (Here, failure to include the positive covariance yields an overestimate of the trajectory of change in marital dissatisfaction for each spouse—results not shown.) More importantly, failure to include the covariance yields an unrealistic model where the marital satisfaction of each spouse supposedly unfolds over time in a manner not affected by the other spouse's marital satisfaction.

Discussion

In this chapter, I have outlined the straightforward extension of REMs and FEMs to LGCMs. Whereas, LGCMs are prevalent in some disciplines such as developmental psychology, they remain rare in family studies. If the phenomenon under study changes across time in a systemic fashion, LGCMs offer the opportunity to model this change. Existing versions of the LGCM are all in the REM framework. That is, the latent terms describing initial levels and change in levels of the dependent variable are assumed to be independent of any other covariates affecting the dependent variable. This is a strong assumption and violations may lead to biased estimates of parameters.

I show that the SEM framework provides researchers a way to test this assumption by allowing estimation of fixed-effect versions of the LGCM. By allowing all or some of the time-varying covariates to covary with the included latent terms, LGC FEM provides a powerful tool, allowing researchers to account for patterns of covariation that may bias parameter estimates.

Further, I show how time-constant covariates can be included in LGCMs. The effects of time-constant covariates may be expressed directly on the latent parameters of the model or directly on the outcome under consideration. Because these two models are nested, a simple LR Chi-square test can be used to adjudicate between the two.

Finally, I show how matched or paired data can be examined within the framework of LGCMs. These models assume that the parameters underlying the trajectory of change across time for one partner affect the same parameters for the other partner. Failure to account for the covariation of these parameters can lead to biased estimated of the underlying parameters.

References

Allison, P. (2005). *Fixed effects regression models for longitudinal data using SAS*. Cary: SAS Institute Press.

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces, 89,* 1–34.

Hjelmborg, J., Fagnani, C., Silventoinen, K., McGue, M., Korkeila, M., Christensen, K., Rissanen, A., & Kaprio, J. (2008). Genetic influences on growth traits of BMI: A longitudinal study of adult twins. *Obesity, 16,* 847–852.

Lyons, K., & Sayer, A. (2005). Longitudinal dyad models in family research. *Journal of Marriage and Family, 67,* 1048–1060.

Neale, M., & McArdle, J. (2000). Structural latent growth curves for twin data. *Twin Research, 3,* 165–177.

Stoel, R. D., van den Wittenboer, G., & Hox, J. (2004). Including time-invariant covariates in the latent growth curve model. *Structural Equation Modeling, 11*(2), 155–167.

Teachman, J., Duncan, G., Yeung, W. J., & Levy, D. (2001). Covariance structure models for fixed and random effects. *Sociological Methods and Research, 30,* 271–288.

Umberson, D., Liu, H., & Powers, D. (2009). Marital status, marital transitions, and body weight. *Journal of Health and Social Behavior, 50,* 327–343.

Veroff, J., Douvan, E., Hatchett, S., & Orbuch, T. (1986–1989). Early Years Of Marriage (Eym) Project (years 1–4). Ann Arbor, MI: University of Michigan, Institute for Social Research, Survey Research Center [producer], 2007. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-10-03. doi:10.3886/ICPSR04557.v1.

Walker, A., Acock, A., Bowman, S., & Li, F. (1996). Amount of care given and caregiving satisfaction: A latent growth curve analysis. *Journal of Gerontology: Psychological Sciences, 51B*, 130–142.

Chapter 2
Families as Coordinated Symbiotic Systems: Making use of Nonlinear Dynamic Models

Nilam Ram, Mariya Shiyko, Erika S. Lunkenheimer, Shawna Doerksen and David Conroy

Families as Coordinated Symbiotic Systems: Making use of Nonlinear Dynamic Models

Families are often conceptualized as continually evolving, relational systems (Cox and Paley 1997; Minuchin 1985). Individual members influence and are influenced by all other members. These reciprocal relations coalesce into family-level symbiotic processes and are the core of study in family systems research and therapy (see Lunkenheimer et al. 2012). Wohlwill (1991) noted that, "... what [reciprocal relationships] would call for are methodologies that allow one to model the interpatterning between two [or more] sets of processes each of which is undergoing change, in part as a function of the other ... The closest approach to this kind of modeling that is indicated for this purpose are probably some of the models from the field of

N. Ram (✉)
Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA
e-mail: nur5@psu.edu

M. Shiyko
Department of Applied Counseling and Educational Psychology, Northeastern University, Boston, MA, USA
e-mail: M.Shiyko@neu.edu

E. S. Lunkenheimer
Department of Human Development & Family Studies, Colorado State University, Fort Collins, CO, USA
e-mail: erika.lunkenheimer@colostate.edu

S. Doerksen
Department of Recreation, Park and Tourism Management, The Pennsylvania State University, University Park, PA, USA
e-mail: sxd49@psu.edu

D. Conroy
Department of Kinesiology, The Pennsylvania State University, University Park, PA, USA
e-mail: dec9@psu.edu

S. M. McHale et al. (eds.), *Emerging Methods in Family Research*, National Symposium on Family Issues 4, DOI 10.1007/978-3-319-01562-0_2,