DISAGGREGATING BETWEEN AND WITHIN-PERSON EFFECTS IN AUTOREGRESSIVE CROSS-LAG MODELS

by

Paul Wesley Scott

B.A., Berea College, 2007

M.A., University of Pittsburgh, 2014

Submitted to the Graduate Faculty of

the School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Paul Wesley Scott

It was defended on

October 2, 2017]

and approved by

Heather Bachman, Ph.D., Associate Professor, Department of Psychology in Education

Clement Stone, Ph.D., Professor, Department of Psychology in Education

Elizabeth Votruba-Drzal, Ph.D., Associate Professor, Department of Psychology

Dissertation Advisor: Feifei Ye, Ph.D., Assistant Professor, Department of Psychology in Education

DISAGGREGATING BETWEEN AND WITHIN-PERSON EFFECTS IN AUTOREGRESSIVE CROSS-LAG MODELS

Paul Wesley Scott, PhD

University of Pittsburgh, 2017

There is often interest in evaluating bidirectional relationships amongst processes over time. Random Intercept Cross-lagged Panel Model (RI-CLPM) and Latent Growth Curve Model with Structured Residuals (LGCM-SR) are two models developed to disentangle within- from between-person effects. These models have shown to out-perform Cross-Lagged Panel Model (CLPM) that confounds between- and within-person effects.

This study uses both empirical and simulated data to compare the performances of these models in assessing the bidirectional relationship between two developmental processes. Data from the Longitudinal Study of American Youth were used to explore bidirectional relationships between student math self-concept and task value from grades 7 to 12. The CLPM indicated Self-Concept dominated Task-Value, while the RI-CLPM and LGCM-SR indicated Task-Value dominated Self-Concept, suggesting that the CLPM's confounding of between- and within-person effects leads to substantively different conclusions than RI-CLPM and LGCM-SR.

A Monte Carlo study was conducted to compare RI-CLPM and LGCM-SR. The RI-CLPM fits time-specific means to capture the functional form of the trajectory, but does not capture variation around the trajectory as a LGCM-SR would. Data were simulated from both models under different causal dominance conditions. For LGCM-SR models, data were generated with different slope variance, covariance, and trajectory shape. Fitting LGCM-SR models to RI-CLPM data results with negative variances of growth factors, suggesting over-parameterization. Fitting RI-CLPM to LGCM-SR data results with underestimation of cross-lagged path coefficients, and the bias is larger in non-dominance conditions and increases with larger slope variances, suggesting the necessity to consider the slope heterogeneity if present. For the nonlinear LGCM-SR data, as RI-CLPM was estimable while linear LGCM-SR always had negative slope variances, capturing the functional form might be more important than capturing variability in slope. Relative fit indices performed well in selecting the correct model between RI-CLPM and LGCM-SR. BIC proved superior at correctly choosing the linear LGCM-SR over the unspecified LGCM-SR.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to express the utmost thanks to my dissertation advisor Dr. Feifei Ye for her invaluable guidance not only throughout the development of this dissertation but throughout my graduate studies.

Further, I would like to thank all of the research methodology faculty who led me through the program and imparted a wealth of knowledge upon me. Dr. Suzanne Lane, who has provided me with immense support and direction over the past several years, Dr. Clement Stone, who generously offered to sit on my committee and provided excellent feedback, Dr. Kevin Kim, who was a great mentor to all and will always be remembered for his enthusiasm, and Dr. Lindsay Page, who gave me wonderful new ways of understanding causal inference.

A special thanks to Dr. Heather Bachman for serving on my committee and providing me with years of research supervision. I am also thankful to have had Dr. Elizabeth Votruba-Drzal on my committee to offer up her unique experience as an applied research to inform the direction of my dissertation.

To all of my friends, family, and colleagues who have been there supporting me throughout my graduate studies. I would especially like to acknowledge my son, Francis whose immense patience with me as I pored over work for the past several years truly made this possible.

1.0 INTRODUCTION

Research within the developmental and educational science requires that we can model longitudinal data in such a way that appropriately capture dynamics over time that are consistent with our theories of change. Such theories of change may entail: (1) what is the functional form of growth, for example, is vocabulary acquisition between 1 to 17 years of age monotonic or do certain spans of time show greater and lesser rates of acquisition; (2) are there differences in development amongst individuals, for example, how does the trajectory of vocabulary acquisition differ between children from higher versus lesser linguistically expressive environments; and, (3) what is the relationship between processes, for example, how does increasing literacy relate to increases in vocabulary acquisition..

In general, we tend to assume that there is both growth within an individual as well as differences in growth between individuals. For example, a child coming from a home where the teaching of math skills was underrepresented will likely enter school with lower math skills than a child coming from a home where mathematic skills were taught prior to school entry, but that child with lower math skills may exhibit an initial rate of growth in math development as a sort of “catching up” process. Even though overall the child from the home with low math representation may exhibit lower math skills than the child from the home with greater math representation, they may exhibit greater growth in math. If our data modeling strategy fails to account for these distinct effects our estimates will fail to give us valuable information about the effects of schooling on early math growth.

In addition to the consideration of separating out between and within-person processes there is also a central consideration of how developmental processes influence one another. Developmental processes do not occur in a vacuum and so it is of interest to delve into the relationships amongst developmental processes. Namely, we may want to know how growth in one domain affects growth in another, do changes in one process cause changes in another, or is there a reciprocated effect. The latter consideration will be central to the subsequent discussion in this paper. Such reciprocation may be exhibited, for example, when considering parenting and child behavior. We may expect parenting behaviors to lead to certain behavioral outcomes in children; however, we would also want to take into consideration the influence that a child's behavior is having on parental responses.

These issues are each addressed with particular canonical approaches: for disentangling between and within-person effects the standard approach has been to use multilevel modeling wherein repeated measures are conceptualized as being nested within individuals. Multilevel models in standard usage can come in either the form of a fixed effect or random effect model. The fixed effect approach will estimate far more parameters because it gives everyone their own intercept term through a set of dummy coded variables, this can also be done with slope terms as well by forming interaction terms between time and the dummy codes. The other, more commonly observed in developmental and educational science, is the random effect model. The random effect model estimates an average intercept or slope and variance components allowing for variation around this average. The variance components account for individual's deviations from the average fixed effect. As one may note the random effect approach requires less parameters than the fixed effect term and so is often favorable in conditions where one doesn't necessarily have the degrees of freedom to spare. Both approaches can be integrated into a latent variable modeling approach and we see more closely how this is accomplished within structural equation modeling in the literature review. Within the developmental and educational sciences such multilevel models for longitudinal data come in the form of what is known as a latent growth curve model (LGCM), which fits the slope representing the fixed slope in the form of a mean value and captures individual deviations through the variance in the latent variable. This approach allows for several functional forms to be captures via path loadings and in some cases fitting additional latent variables. The evaluation of bivariate relations across time is typically addressed by using a cross-lag panel model (CLPM), such models autoregress the repeated measures of a construct then predict the measures on the other construct across lags. When using such models, we are generally looking for a causal process, specifically we wonder if one process is dominating another or if we have a fully reciprocated model. Examples of these considerations of bivariate processes can be seen in phenomena such as when parental behaviors influence child behavior and vice versa. Recent attempts have sought to merge these two methods in the hopes of gathering information on processes theorized to have both bivariate relations and to contain a between and within person component..

1.1 STATEMENT OF THE PROBLEM

The cross-lagged panel model (CLPM), also known as the Autoregressive Cross-Lagged (ARCL) Model, has become a standard approach for modeling the relationship between two or more developmental processes which are believed to have a mutual influence on one another.

These models take repeated measures from multiple constructs, generally bivariate, and then fit autoregressive paths across the lags within the construct measures and cross-lag paths between constructs. The spirit of such a modeling approach is that by accounting for the influence of a construct on itself over time, we can better gauge the influence of a construct on another construct over time (Campbell, 1963; Kenny, 1973). One of the primary interests in adopting such models is to test for causal mechanisms. Specifically, we wonder if one process causally dominates another or if the two processes are reciprocal to one another, i.e. no one process is dominating the other but both are mutually influential on one another. For example, we may wonder the nature of how parental behavior is influencing child behavior or vice versa. Another area where interest in such models have abounded is in the examination of mediation mechanisms (Selig & Preacher, 2009; Maxwell & Cole, 2007) since such mechanisms are definitively causal in nature, unveiling over time, and involving multiple constructs. Therefore, the value of cross-lag panel models is apparent for this purpose.

The interpretation of the paths pertains to changes in rank-ordering of individuals over time. Since such an interpretation is limited to the between person level, the cross-lagged panel model has been criticized for not providing information that is most relevant for developmental researchers where the questions are often concerned with changes that happen at a within individual or within dyad level. This issue of disaggregating the between and within person models has been well discussed within the multilevel and structural equation literature (e.g., Curran & Bauer, 2011; Wang & Maxwell, 2015; McArdle & Epstein, 1987; Preacher, 2008). It has become increasingly common that we fit these multilevel models in the form of a latent growth curve model (LGCM) because of the flexibility in specifying such models. The basic approach is to fit the model with an intercept and slope which have a fixed and random component. The fixed components are represented by mean values and individual deviations are captured through the variance of the latent variables (termed intercept and slope factors in LGCM). In this way, we are allowing everyone to have their own specific trajectory. This approach also allows for a growth trajectory to take various forms via path loadings or fitting additional latent variables.

To focus our cross-lag models on the level of inference which is often of most interest in developmental research, namely the within person/dyad level, recent innovations have sought to bring the merits of the growth curve approach to the cross-lagged panel model. These models share the common feature of fitting latent variables to panel data. Hamaker, Kuiper, and Grasman (2015) proposed a cross-lag panel model that utilizes a random intercept to separate out stable between person differences and fits time specific means to allow for overall group changes across time (RI-CLPM). The cross-lag portion is modeled on the residuals such that our examination of the bivariate processes influences on one another concerns the time specific individual deviations across given time points. Though this model does capture some sense of the overall change in constructs across people over time, they do not account for individual variation in growth trajectories. Curran, Howard, Bainter, Lane, & McGinley (2013) similarly proposed utilizing the residuals in such a way, except in their models they also fit a random slope term that allows for individual variation in growth over time. They refer to this model as the latent growth curve with structured residuals (LGCM-SR). The two models take the same basic approach to disaggregating effects in cross-lagged panel models, differing only in how they account for change over time. The RI-CLPM with time specific means doesn't account for variation amongst individuals over time. However, it does not require a specification of the overall functional form of growth as the LGCM would, since the time specific means are freely estimated without regard to the means at other time points. If the means were constant across time, then the intercept would represent this as its mean with time specific means would constrained at zero. This would essentially represent an LGCM-SR where the slope parameters were all zeroed out. The issue of non-linearity in the mean trajectories of a construct over time can theoretically be accommodated in each of these models. However, the latent growth curve with structured residuals will require additional specifications to the slope parameters, meaning the researcher will need to have some insight into the functional form of a developmental process.

The above growth curve approaches were developed as random effects models. There are some related, fixed effect versions of autoregressive cross-lag models currently being developed (Williams, Allison, & Moral-Benito, 2015; Allison, 2005; 2015; Allison, Williams, & Moral-Benito, 2017). The motivation underlying such models is that the inclusion of lagged dependent variables is problematic for the assumption of independence between error terms and predictors, what is known in the econometric literature as the “incidental parameter problem”. Additionally, there is a concern with how initial conditions should be handled. Random effects models do not adequately address these problems. The above-mentioned methodologists have been in the process of developing fixed-effects models which can incorporate lagged dependent variables. Due to the complexity of specifying and estimating such models, these haven’t yet been fully adopted into the familiar reciprocated model represented in the standard CLPM formulation. Though the authors (Allison et al., 2017) indicate that such a specification can be implemented in structural equation modeling software, no such model is demonstrated[1]. The current versions of

[1] The models demonstrated in Allison et al. (2017) specify that the intercept fit to the dependent variable series correlates with the time varying predictors from which the cross-lags originate. In the case where we have a reciprocal model, then those time varying predictors will now be lagged and have an intercept, thus will be

these models involve analyzing one dependent variable series at a time, i.e. one requires separate models to check the cross-lags from X to Y versus the cross-lags from Y to X. For this reason, our evaluation of fixed effect cross-lag models will be limited as they are not directly comparable to the RI-CLPM and LGCM-SR specifications.

Few studies exist evaluating the performance of the above mentioned multilevel cross-lagged panel models. Hamaker et al. (2015) did an empirical and simulation study comparing the CLPM to the RI-CLPM they developed. Because their study was only focused on controlling for trait-like stability in constructs over time, no consideration was given to either change in the construct over time nor variation in such change. Berry and Willoughby (2016) demonstrated via simulation and empirical examples how the cross-lagged panel model may be misrepresenting the relationship between parental discipline and children's aggressive behaviors, and suggest using the LGCM-SR as an alternative. In their study, the growth trajectory was only set to capture mean, linear change ignoring variance in slopes amongst individuals. Curran et al. (2014) in formulating the LGCM-SR give empirical examples of how to use the model. However, they did not conduct a genuine simulation study comparing performance under various conditions. In this study, we hope to fill in these gaps.

1.2 PURPOSE OF THE STUDY AND RESEARCH QUESTIONS

In this study our primary focus will be on the random intercept cross-lag panel model and the latent growth curve model with structured residuals as these are currently established in such a

endogenous. This specification seems problematic to the basic aim of this model, viz. to prevent dependence between the predicting values and the error terms, since simultaneous error terms and predictors aren't cleanly separated.

way as to allow us to disentangle between and within-person effects when evaluating reciprocated relationships to determine causal dominance. Because these models are accounting for between person differences and the cross-lagged portion is concerning the within-person processes, it is of interest to explore how variations in the between person components affect the conclusions we derive about the within-person processes, especially in relation to change over time. For example, since the RI-CLPM model does not account for between person variation in change, it is of interest to see what the consequences of this omission will be.

The purpose underlying this study is to explore, in a comparative manner, models that have been proposed to evaluate reciprocated relations between developmental processes while disentangling between and within-person effects. Further, we want to determine when one model becomes preferable to another. Being relatively recent in their development these models have not been widely applied or studied. In the following we will present the models as fitted to real data to show how our conclusions about the relationship between developmental processes can vary due to model selection. For a more formal evaluation we will conduct a simulation study to gain insight about how specifications of the between-person differences in change and the functional form of a trajectory can influence the conclusions we derive about the cross-lag relations. The hope is that in the process of exploring these models researchers can gain a better sense of how to use these models, and under what circumstances one model becomes preferable to another.

These issues will be more closely explored in the following sections where we first review at some depth the relevant literature, and then give an empirical demonstration of the models followed by a simulation study to compare and evaluate the models in terms of how the slopes' functional form, variance and covariance influence a model's fit and the conclusions we derive about the cross-lag relations. Our general research questions will pertain to

(1) how these models will bring us to different conclusions depending on what types of models we fit to empirical data;

(2) how are different criteria of model fit influenced across the conditions and models being fit to data;

(3) how accurate are our conclusions about causal dominance under conditions of varying magnitudes of variability and covariability in the slopes;

(4) how will the functional form of a growth trajectory influence the relative performance of RI-CLPM versus a Linear LGCM-SR.

Namely, we postulate that as variability and covariability in the slopes increases the RI-CLPM will lead to more misrepresentative cross-lag parameters and hence negatively impact the conclusions about causal dominance we derive from it. In terms of functional form, non-linear trajectories will degrade the linear LGCM-SR potentially giving the RI-CLPM an advantage in properly representing change over time. However, this improvement will likely be offset by the amount of slope variability.

1.3 SIGNIFICANCE OF THE STUDY

As mentioned above these models are relatively recent in their development and hence have not been widely utilized in research. One part of this research then is to hopefully familiarize developmental researchers of both the availability of such models as well as the capabilities of these models for given research purposes. A further consideration is that there hasn't been much methodological evaluation of these models in relation to one another. In fact, Hamaker et al. (2015) make no mention of the Curran et al. (2014) model which is, as noted, related to the model they propose. Curran et al. (2014) did not provide a simulation study to show the relative methodological merits of their proposed model.

The importance of this study is in its value to applied research in the developmental sciences where the theory of change of interest may require models that disentangle between person differences from within person processes where developmental process have a reciprocal relationship across time. Often, developmental researchers are mainly interested in how developmental processes at the intra-individual level are influencing one another, making the historic lack of models for teasing this out a major limitation. In this study, we will introduce some potential models that can help in this task. Furthermore, we will be evaluating the conditions where the different models will perform best.

Additionally, the value of such models to mediation analysis should be noted. Cross-lag panel models are one of the most common and trusted methods for testing posited mediation processes. However, CLPM is limited in that it only accounts for changes in rank ordering without separating the within and between person levels. These novel modeling approaches will especially benefit researchers concerned with mediation processes.

Overall, this study serves as an overview and evaluation of models which allow researchers to gather specific information about how developmental processes unfold over time and influence one another at the individual level. Given the power of these models in capturing information that is more aligned with our theories of development, their value is immediately clear. However, before we apply these models with confidence we should have some sense of how they perform under varying circumstances. This is the issue explored in this paper, and the findings we gather will serve to inform developmental researchers in the application of these models.

2.0 LITERATURE REVIEW

In the following we will consider in turn different modeling strategies for addressing inferences concerning ARCL relations and the disentanglement of between and within person effects in longitudinal data as we work our way towards modeling strategies that subsume these considerations together.

2.1 CROSS-LAGGED PANEL MODEL

The most well utilized modeling approach for exploring bidirectional effects amongst processes is the Cross-Lagged Panel Model (CLPM). The strength of the CLPM is that it allows for a close examination of reciprocated effects while controlling for the stability in each construct. Thus, when researchers have questions about bidirectional relations between constructs over time the CLPM becomes a desirable solution. The spirit of the CLPM is founded in questions of causal processes. Hence, we often find this model crop up in mediational research using repeated measures (Cole & Maxwell, 2003). Since mediation is proposed as exploring causal mechanisms the CLPM is perfectly suited because of its ability to disentangle directionality in effects. Some mediation processes that have been explored is the extent to which perception of criticism mediates the influence of maternal criticism on adolescent anxiety and depression (Nelemans et al., 2014); also, how early development in self-regulatory behaviors mediate the link between development of language ability and the development of inattentive-hyperactive problem behaviors (Peterson et al., 2015). A more general use of CLPM in developmental research is exploring the extent to which factors involved in a developmental context are influencing one another in a causal manner. Specifically, we want to explore the extent to which one process dominates another in such a way that its influence does more in effecting change in the other variable, instead of the reverse. Moreover, we may also have what is known as a reciprocated relation, wherein changes in one process lead to changes in the other and vice versa, but neither is regarded as having a dominant effect. These are bivariate models, and it is what we will most closely consider in this paper. Research in this mode has found perceived loneliness causally dominating depressive symptoms in 50-68 year old individuals (Cacioppo et al., 2010); MacKinnon (2012) found that increasing academic achievement led to increasing perceptions of social support amongst kids transitioning into college; there is also a wealth of research exploring the reciprocated effects of parent and child behaviors (e.g., Shaffer et al., 2013; Burke et al., 2008; Pardini et al., 2008).

CLPM allows this close examination of directionality amongst processes by closely controlling for the covariances amongst repeated measures of two variables (in the bivariate case). To discuss the CLPM let's propose two hypothetical variables, X & Y. In the following we will assume there is no third variable intervening with the with the relationship between X & Y. There are now 4 possible causal conditions that can exist between X & Y, neither X causes Y nor does Y cause X, X causes Y but Y does not cause X, Y causes X but X does not cause Y, or, both X causes Y and Y causes X. The beauty of having repeated measures is that we can evaluate such causal conditions. With cross-sectional data, all that can really be concluded is that X and Y are related, and we must postulate the causal precedence we believe to be in action either through a theoretic justification or by indicating the temporal precedence of what a measure reflects. For example, if we acquired math scores at grade 7 and we had information about the type of elementary score attended by a child we have some justification in assuming that elementary school features causally precede the middle school math scores. With the cross-lagged panel model we can test for these causal relations.

The three main components that characterize the CLPM are the autoregressive portion, the cross-lagged portion, and simultaneous correlations. Rogosa (1980) indicated when comparing cross-lag pathways we must account for stability in processes over time (in the form of autoregression) to derive the correct causal conclusions, because if two processes are characterized by varying degrees of stability we can acquire spurious correlations which will confound the comparison of cross-lag pathways. The essential motivation for autoregression comes from the fact that there is dependence amongst repeated measures of a variable. Often, we find that prior measures are the best predictor of measurement at subsequent time points, thus we model this in the CLPM as an indicator of how stable a process is over time. In the most basic sense we establish a single autoregressive path in the lag from t to t-1. In a sense this would be considered a first order autoregressive model; however, in some cases people are motivated to add additional orders such that there is a path for the t to t-1 as well as for the t to t-2 lag. We can consider this a second order autoregressive model. Additional paths can be added for various lags. The purpose for entering these autoregressive paths is to first control out variation in a measure that can be explained by prior measures so that we can get a more accurate estimate of the influence between X and Y. To adequately model the influences between X and Y, we begin by taking the exogenous X and Y variables from t=1 and account for their initial, simultaneous relationship between one another by allowing them to covary with one another. The residuals resulting from autoregression at subsequent time points can then be correlated to further account for any variation that can be attributed to the simultaneous relationship between X and Y. In this way, we can focus more closely on the parameters of primary interest for evaluating the causal condition that represents the relationship between X and Y, namely our cross-lagged pathways. Having accounted for the dependence of a process on itself over time as well as the simultaneous relations between two process we can now allow one variable from prior time points to explain out any remaining variation in the other variable. Then we can conclude that one variable is predicting another variable at subsequent time points above and beyond what is explained through autoregression and contemporaneous bivariate correlation. In this way, we can see conceptually how the CLPM serves as a powerful model for evaluating causal relations amongst processes by closely controlling sources of errors in repeated measures of multiple variables. Once we have these cross-lagged estimates we can evaluate them to determine the causal conditions characterizing bivariate processes of development. This comparison is done on standardized coefficients in order to ensure that the scale of each variable is not driving the interpretation (Bentler & Spackart, 1981).

We can represent CLPM for 3 time points diagrammatically as in the left hand portion of Figure 2. This is the representation which we will be sticking with throughout this paper, with two processes an autoregressive path, an X->Y cross lag path, and an Y->X for each t to t-1 lag, along with a correlation between variables at each time t. The CLPM formulation we will be working with is put forward in Hamaker et al. (2015). The CLPM gauges changes in rank ordering over time, thus time specific means are fit at each time point and we enter individual deviations from those means into the Autoregressive Cross-Lag structure for the model. We decompose the individual score at each time point as such:

$$x_{it} = \mu_{x_t} + \varepsilon_{x_{it}} \qquad (1a)$$

$$y_{it} = \mu_{y_t} + \varepsilon_{y_{it}} \qquad (1b)$$

Where the respective μ correspond to the time specific means for the given variables and the ε indicate the i^{th} individual's time t^{th} specific deviations from the mean. We use this approach to allow us to structure the residuals into the autoregressive cross-lag structure of the model for reasons that will become apparent as we move along. We write this portion of the model as such (for single lag auto-regression and cross-lag):

$$\varepsilon_{x_{it}} = \rho_{xx}\varepsilon_{x_{it-1}} + \rho_{xy}\varepsilon_{y_{it-1}} + v_{x_{it}} \qquad (1c)$$

$$\varepsilon_{y_{it}} = \rho_{yy}\varepsilon_{y_{it-1}} + \rho_{yx}\varepsilon_{x_{it-1}} + v_{y_{it}} \qquad (1d)$$

where the first term on the right hand side of the equations (1c-d) represents the autoregression of the individual's time specific deviation at time t on the individual time specific deviation at the immediately prior time point, the second term expresses the cross-regression of a given individual's time specific deviation for one variable on their time specific deviation on the other variable, and the final term now takes the place of residual error variance for a variable at time t for individual i . Substituting through gives us the composite model decomposing each of our X and Y values at a given time as:

$$x_{it} = \mu_{x_t} + \rho_{xx}\varepsilon_{x_{it-1}} + \rho_{xy}\varepsilon_{y_{it-1}} + v_{x_{it}} \qquad (1e)$$

$$y_{it} = \mu_{y_t} + \rho_{yy}\varepsilon_{y_{it-1}} + \rho_{yx}\varepsilon_{x_{it-1}} + v_{y_{it}} \qquad (1f)$$

with residual covariance matrix:

$$\psi = \begin{bmatrix} \sigma^2_{x_{t-s}} & & & \\ \sigma_{x_{t-s},y_{t-s}} & \sigma^2_{y_{t-s}} & & \\ 0 & 0 & \sigma^2_{v_{x_t}} & \\ 0 & 0 & \sigma_{v_{x_t},v_{y_t}} & \sigma^2_{v_{y_t}} \end{bmatrix}$$

where t-s indicates, the earliest given time point in the series (t=1).

The autoregressive parameters, denoted with ρ_{xx} & ρ_{yy} measure the stability in a

construct over time thus useful for allowing to have a sense of how much growth there is in a process over time for individuals, as the magnitude of this parameter estimate increases the more stable an individual's status on that given variable is over time. Hypothetically, a value of one would represent a perfect prediction of a variable from its prior values implying that relative to others an individual status on some construct stays the same from one time to the next. This specification focuses on the structured residuals which are reflecting the time specific deviations for an individual, suggesting that what we are evaluating in the cross lagged panel model is changes in rank ordering of individuals in relation to other. If we were to fit the cross-lagged panel model to the observed variables the results would essentially be the same.

By fitting the time specific means and evaluating the individual deviations from this we are simply rescaling the variable such that we can consider individual scores as being centered about the occasion specific mean. Hence, the focus is on the inter-individual differences, i.e., rank ordering from one occasion to the next. Stability in this sense is interpreted as how consistent the relative standing amongst individuals is across time. However, it doesn't capture how much an individual stays constant or changes on the variable over time. In other words, it doesn't capture the within-person growth process. This shortcoming of the cross-lagged panel model is one of the most commonly cited shortcomings of its application. This motivates attempts to establish a technique for disentangling the within person level from the between person level. An immediately logical place to consider such modeling techniques is in the realm of multilevel modeling.

2.2 MULTILEVEL MODELING WITH LAGGED DEPENDENT VARIABLES

Multilevel modeling is explicitly designed to separate levels of analysis from one another. The basic approach is to person mean center a variable at the repeated measures level and enter the individual's mean at the individual level. Readily one can see that we are limited by the fact that the estimation of a multilevel model for repeated measures requires the repeated measures to be nested within individuals such that the repeated measures are structured as univariate and correlated within the individuals. After doing this, the estimation of a time series based in predicting from lag to lag is prevented. An outcome being nested in an individual is the variance which we are explaining with the predictors and is assumed to be uncorrelated with the values on the predictors. Namely this is the assumption that the predictor variables are uncorrelated with the residuals. Under the framework where a lagged dependent variable is a separate variable from the outcome variable this is of no great concern. However, when the lagged dependent variable is included with the outcome we have a violation of this assumption (Allison, 2015). We already know that the lagged dependent variable relates directly to the error as it was used in the estimation of the random components.

To illustrate, consider the random intercept model. We have some time varying dependent variable, y_{it}, for *t* time points clustered in each individual *i*, that we model to have a fixed value for the intercept β_0 that we give a random distribution to by allowing everyone to have their own unique deviation, u_{0i}, from this grand mean intercept such that we have an intercept distribution, $u_{0i} \sim N(0,\sigma_u^2)$. This random intercept term conceptually captures the influence of individual time-invariant unobservables on the given values of y across time not captured by the other explanatory variables in the model. Including some time varying predictor x_{it} and the lagged dependent variables $y_{i(t-1)}$, the model can be represented as,

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 y_{i(t-1)} + u_{0i} + \varepsilon_{it} \qquad (2),$$

As is usual we assume that the error terms u and ε will be independent of the other explanatory variables since the error terms are representing what is left unexplained (i.e. unobservables) by the observables (i.e. the explanatory variables). Herein lies the heart of the problem when fitting the lagged dependent variables in the random effect model, the lagged dependent variables cannot be assumed independent of the random component.

Due to this Allison (2015) suggests fitting a fixed effect model (Allison, 2005; Bollen & Brand, 2010). The basic idea underlying these fixed effects models is that instead of creating a random component parameter we instead create a set of dummy variables such that each individual, aside from some individual who will serve as a reference, will have a variable d=1 on any observation collected from that individual and d=0 otherwise, the reference individual will have d=0 across all observation for all individuals including their own, such that every other individuals' fixed effect is interpreted in reference to this reference individual. Without this reference individual, the model would be overparameterized. The fundamental conceptual distinction between the random and fixed-effects models is that in fixed effect models we are given every individual their own intercept, whereas with random effects we are simply distributing individual intercepts around some average intercept. Note here: we can also fit slopes as such, namely we can give a distribution to slope terms (random effects) or give everyone their own slope value (fixed effects). We can write this basic distinction for the intercept only models as such:

$$y_{it} = \beta_{i0} + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad (fixed\ effect) \qquad \text{(3a)}$$

$$y_{it} = \beta_0 + u_i + \varepsilon_{it} \qquad u_i \sim N(0, \sigma_u^2) \quad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad (random\ effect) \qquad \text{(3b)}$$

As can be seen above the fixed effect model denotes that each individual *i* are represented by a unique intercept, β_{i0}, and only the residual variance has a distribution, whereas in the random effect model we have a fixed intercept and individual deviations, $\beta_0 + u_i$, while both the time-level residual (ε) and the intercept residual (u) are randomly distributed. Within this fixed effect framework, the intercept parameters, while accounting for clustering, are not assumed to be uncorrelated with the other explanatory variables. Entering lagged dependent variables into this model no longer poses a threat. Fitting this model within a OLS type framework still requires that we establish our outcome as univariate, while lagged DV is entered as another variable in the data. Both the fixed and random effects models can be considered as cases of multilevel models, in so far as we are nesting observations within individuals. The intercept that accounts for clustering can be viewed as a latent variable, in so far that in both models the function of the intercept term is to account for time invariant unobserved factors affecting the individual's outcome observations across time (Bollen & Brand, 2010; Allison, 2005). Upon fitting these intercept factors as latent variables, the only major distinction is that within the random effect framework the intercept term is uncorrelated with the predictor variables while the fixed effect model correlates the intercept term with the predictor variables. Figure 1 gives a general representation for distinguishing the models within a structural equation format. The dynamic linear panel model, i.e. fixed-effects models with lagged dependent variables (Bollen & Brand, 2010; Williams, Allison, & Moral-Benito, 2015; Allison, Williams, & Moral-Benito, 2017; Allison, 2005) is a currently developing model which may in the future offer a fixed-intercept option for the cross-lagged model that can evaluate reciprocated cross-lag pathways in a bivariate model while controlling out time-invariant person level fixed effects. However, in its current state this has yet to be fully established. It is tricky to relate the two processes in a clean way, both in terms of conceptualizing the model as well as computational execution of such a model. However, this does not entirely prevent us from gathering some information concerning causal dominance. We will simply be required to fit two separate fixed-intercept models and compare cross-lag coefficients from the separate models, e.g. one for testing the effects of X on Y with Y having lagged DVs and another for testing the effects of Y on X with X having lagged DVs:

$$y_{it} = \mu_t + \rho_{yx} x_{i(t-1)} + \rho_{yy} y_{i(t-1)} + \alpha_{y_i} + \varepsilon_{y_{it}} \tag{3d}$$

$$x_{it} = \mu_t + \rho_{xy} y_{i(t-1)} + \rho_{xx} x_{i(t-1)} + \alpha_{x_i} + \varepsilon_{x_{it}} \tag{3e}$$

The equations above represent the fixed-intercept model, which will be estimated separately, so we consider each formula as representing a separate series. We have ρ_{xx} & ρ_{yy} representing the autoregressive coefficients for the lagged dependent variables, ρ_{yx} & ρ_{xy}, represent the cross-lag coefficients, with the α terms representing the fixed-intercept for individuals, and ε terms for time specific errors for the i^{th} individual and time t. All of the predicting elements are correlated as they are considered exogenous, and within each series the error, ε, at each time point must correlate with future values of the time-varying covariate, i.e.

$Cov\left(\varepsilon_{x_{it}}, y_{t+s}\right) \neq 0 \quad and\ Cov\left(\varepsilon_{y_{it}}, x_{t+s}\right) \neq 0$, where t+s represents measurement occasions of a variable up to s lags into the future. This assumption is key to gaining an estimate for reciprocal effects since it allows for the influence of one variable at prior times on later realizations of the other variable be accounted for. We will see a demonstration of this model at a later point in this paper

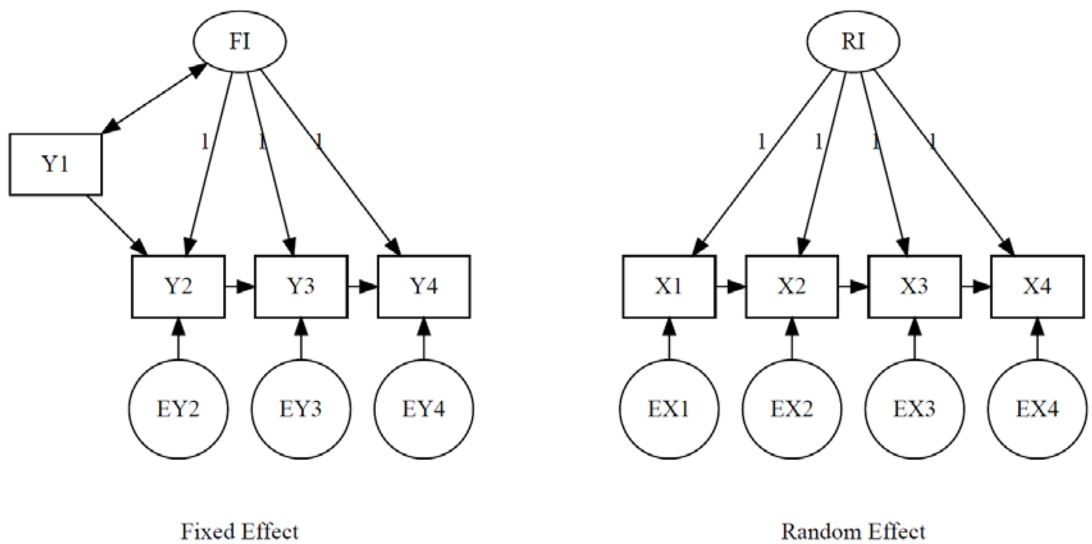


Figure 1. General Panel Models for Fixed and Random Effects containing Lagged DVs

2.3 RANDOM INTERCEPT CROSS-LAGGED PANEL MODEL

Random Intercept Cross-lagged Panel Model is a more recently developed model to address the reciprocated effects while disaggregating the within-person effect from the between-person effect (i.e., stability). Hamaker, Kuiper, and Grasman (2015) proposed a multilevel model for assessing cross-lag parameters wherein a random intercept is fit to a CLPM (RI-CLPM). However, they state that this random intercept (which is fit in the form of a latent variable) is rather a representation of a person's trait like stability in a construct over time. They begin in stating their motivation for doing so is that we are often not controlling for the right kind of stability when fitting standard CLPM. As mentioned before the CLPM controls for temporal stability in a process over time, which is to say that everyone is varying around the same means across time. Thus, there is no accounting for stable trait-like differences between individuals.

By the inclusion of the random intercept we now control for this trait-like stability to disentangle the between and within person levels of analysis. The controlling for trait-like stability can be understood as an omitted variable problem, wherein we are accounting for unobserved time-invariant characteristics influencing the estimation of the cross-lag pathways. The partialing out of the between-person variance via inclusion of the random intercept results in the interpretation of the cross-lagged parameters as referring to the within-person processes. Often this is the main interest of research in developmental sciences, where the level of inference is on the nature of individual development. Thus, there is clear motivation for pulling out the inter-individual differences that endure as stable over time.

This modified model can be understood as now controlling for both temporal stability (i.e. how stable an individual stays on a construct from one time to the next) as well as trait stability (i.e. the varying degrees to which different individuals stay more or less stable on a trait across time). Figure 2 gives a portrayal of the RI-CLPM in comparison with CLPM.

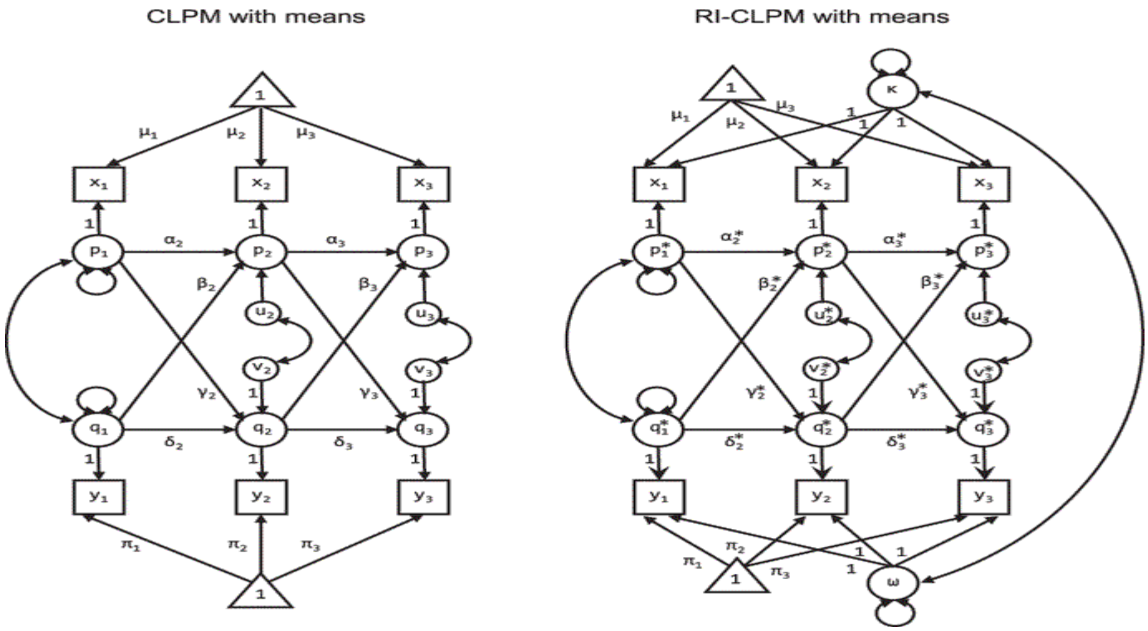


Figure 2. CLPM and RI-CLPM extracted from Hamaker et al. (2015)

The RI-CLPM is simply an extension of the CLPM which decomposes the score as such:

$$x_{it} = \mu_{x_t} + \alpha_{x_i} + \varepsilon^{*}{}_{x_{it}} \qquad (4a)$$

$$y_{it} = \mu_{y_t} + \alpha_{y_i} + \varepsilon^{*}{}_{y_{it}} \qquad (4b)$$

As was the case in the CLPM model μ represents the time specific means and ε are individual deviations from the time specific means, the additional term, α, are the intercepts which capture the trait like deviations from these means. From this formulation, we can appreciate that we have now included a person-mean centering of the variables in the model, which is akin to the standard multilevel approach for disentangling between and within person effects wherein we center individual scores around the individual's mean score across time points. In fact, we would expect a high similarity in results from a standard CLPM performed on person-mean centered variables as would be obtained from the RI-CLPM. This will be examined in a later section. The following cross-lag and autoregressive formulations will look highly similar to in 1c-d, so we will distinguish these by including an asterisk (*) since the estimates & interpretation will be different. The estimates will now have the trait stability being accounted for, thus giving the interpretation that the deviation terms now represent individual deviation from the time specific group mean as well as the deviation from their own mean across time, $\varepsilon = x_{it} - (\mu_t + \alpha_i)$. In this way, we see that we have now centered around both the time-specific and person-specific means, and so all corresponding coefficients will be interpreted in this light.

$$\varepsilon^*{}_{x_{it}} = \rho^*{}_{xx}\varepsilon^*{}_{x_{it-1}} + \rho^*{}_{xy}\varepsilon^*{}_{y_{it-1}} + v^*{}_{x_{it}} \qquad (4c)$$

$$\varepsilon^*{}_{y_{it}} = \rho^*{}_{xx}\varepsilon^*{}_{y_{it-1}} + \rho^*{}_{yx}\varepsilon^*{}_{x_{it-1}} + v^*{}_{y_{it}} \qquad (4d)$$

Substituting through we show the completed model as:

$$x_{it} = \mu_{x_t} + \alpha_{x_i} + \rho^*{}_{xx}\varepsilon^*{}_{x_{it-1}} + \rho^*{}_{xy}\varepsilon^*{}_{y_{it-1}} + v^*{}_{x_{it}} \qquad (4e)$$

$$y_{it} = \mu_{y_t} + \alpha_{y_i} + \rho^*{}_{xx}\varepsilon^*{}_{y_{it-1}} + \rho^*{}_{yx}\varepsilon^*{}_{x_{it-1}} + v^*{}_{y_{it}} \qquad (4f)$$

with residual covariance matrix:

$$\psi = \begin{bmatrix} \sigma^2_{\varepsilon^*_{x_{t-s}}} & & & & & \\ \sigma_{\varepsilon^*_{x_{t-s}},\varepsilon^*_{y_{t-s}}} & \sigma^2_{\varepsilon^*_{y_{t-s}}} & & & & \\ 0 & 0 & \sigma^2_{v^*{}_{x_t}} & & & \\ 0 & 0 & \sigma_{v^*{}_{x_t},v^*{}_{y_t}} & \sigma^2_{v^*{}_{y_t}} & & \\ 0 & 0 & & & \sigma^2_{\alpha_x} & \\ 0 & 0 & 0 & 0 & \sigma_{\alpha_x,\alpha_y} & \sigma^2_{\alpha_y} \end{bmatrix}$$

The autoregressive parameters no longer represent stability in rank ordering over time but rather a kind of "carry-over" effect within an individual from one to the next. Namely, the parameters are interpreted in relation to an individual relative standing based on their own expected score (i.e., $\mu_t + \alpha_i$), for example, a positive autoregressive coefficient implies that when an individual scores above their expected score at one time point they will also score above their expected score at the subsequent time point. Similarly, the cross-lags represent the degree to which an individual's deviation on one variable predicts the deviations on the other variable. We are now evaluating the cross-lag parameters as difference scores controlling for trait stability, so the dynamics between the processes is being assessed at the within-person level.

The relationship between the standardized cross-lag pathways from the CLPM and RI-CLPM is a function of the within person cross-lag coefficient, autoregressive coefficient, the covariance of the within person deviations at t-1, the variance of the within-person deviation at t-1, the variance of the between person trait-like stability for the construct, and the covariance between the random intercepts for the different constructs. Thus, the degree, direction, and way in which the cross-lag paths will differ between the CLPM and RI-CLPM will depend on these factors (Hamaker et al., 2015). Furthermore, as Schuurman et al. (2016) indicate the way in which we standardize the coefficients will also bear upon the conclusions that we derive about cross-lag associations, potentially changing the causal conclusions we derive and the size of the effect we observed.

The type of standardization that we use depends on what component of variance we utilize, i.e. total variance, within-person variance, or between-person variance. These considerations have been explored in terms of the more general topic of effect sizes in multilevel models (Hedges, 2007), where the question of which variance component you are using in standardization is a matter of what sort of effect size interpretation you want, or in other words, to which level of analysis are you wanting to evaluate. The within-person standardization will be calculated for everyone as how much variance they exhibit on a variable across time. We can then take the unstandardized coefficient and multiply it by the ratio of the respective variable variance, where the standard deviation for the predictor variable is in the numerator and the standard deviation for outcome is in the denominator. This ratio is gauging how variance in the predictor is explained by the variance in outcome, hence the standardized path will be largest for the predictor variable that varies the most in an individual. These paths are now person specific, so to get a more general fixed effect that gauges paths across the within-person levels, we will get a pooled average of all the person specific coefficients. This type of standardization is interpreted as the amount of standard deviations that an outcome will change for each increase in standard deviation in the predictor specific to the individual level, thus the fixed effect says we expect the changes in a prior measure on a predictor to predict a corresponding amount of change in the outcome at the subsequent time point for any given individual.

The between person standardization simply uses the variance in the person-specific means across time. The logic for deriving the standardized coefficient is the same as with the within-person standardization except in this case we are using the respective between-person variance of means in the ratio resulting in a coefficient which now represents the amount of change in standard deviation in individual's means in the outcome we predict from a standard deviation change in the person-specific mean on the predictor. For the fixed effect, we take the expected fixed effect as calculated for the within-person level (note: the cross-lags are at the within-level) and cast them in terms of the ratio as used for the between person coefficient estimates, thus we are now saying we expect an individual's prediction of the outcome to correspond to a change in individual mean standard deviations for a standard deviation unit change in the individual predictor means. Thus, the evaluation is in terms of person specific means, which process is causally dominant.

The grand standardization combines both the within and between-person level variances. This now cast the analysis in terms of grand standard deviations, and can almost be seen as a population level analysis as it is polling across both variance in person-specific means as well as the time-specific variance within individuals. The calculation of the coefficients and fixed effects is that same as was done for between person standardization, but we also include the pooled within person variance in the ratio. In most cases, it would seem that the within-person standardization is preferred since when examining developmental processes, these are occurring at the within-person level and thus the intra-individual changes seem most relevant. However, different research may warrant different levels of interest, but it should be noted that when using between person and grand standardization you will still be basing conclusions off the within person cross-lag parameter. Given all this one thing that can be concluded is that each kind of standardization will be yielding different numeric values. Schuurman et al. (2016) state that it appears that MPlus uses within person standardization, though they are not clear on how this is being achieved. In our studies the paths can be interpreted as within person standardized as we will be estimating all our models in MPlus and using the standardized coefficients for interpreting causal dominance.

The inclusion of the random intercept will require that we have at least 3 waves of data to identify the model, whereas the CLPM will only require two waves. As is usual with structural equation models we can increase our degrees of freedom by fixing parameters. This is of value providing that the fixing is reasonable in its assumptions; for example, if we have equal intervals it may be reasonable to assume that the influence of a variable on itself and on the other variable will be consistent across lags, thus allowing us to fix the autoregression in X and Y (yielding 2 more degrees of freedom) and the cross-lags between X-->Y and Y-->X across lags (yielding an additional 2 degrees of freedom), for a total of 4 additional degrees of freedom (df=5) for models containing 3 waves. A more tenuous constraint would be to assume that the means are also consistent, but there are many cases where this would not be the case when considering developmental processes.

When we anticipate structural changes within a process over time we may further relax loadings on the intercept. However, in doing so we change the essence of the model; we no longer control trait stability through a random intercept but rather we are simply assessing a trait in the more common sense, which may be able to vary over time. Alternatively, as we will see later we could add a growth curve, which would allow for us to both control for time-invariant trait stability as well as trait change. These respective types of models are in the family of what can be referred to as latent state-trait models (e.g., Luhmann et al., 2011; Steyer et al., 2015) and latent growth curves with structured residuals (Curran et al., 2014).

2.4 RELATED SEM APPROACHES

Different structural equation models do similar things as the RI-CLPM, namely the aim is to control for between person, trait like stability. In the following we will briefly consider these related models and discuss their relationship to the RI-CLPM, which may give researchers some indication as to which model fits their research and data best. The first model to mention is the Latent Growth Curve with Structured Residuals (LGCM-SR; Curran et al., 2014; Curran & Bollen, 2001), which we will consider in more depth later. For now, it is worth mentioning that this model is essentially the same as the RI-CLPM but adds a latent slope term. Thus, the LGCM-SR reduces to the RI-CLPM when the loading on the slope terms are equal to zero. The LGCM-SR becomes preferred in cases where we anticipate that we are working with a process that demonstrate a considerable growth trajectory that varies amongst individuals, i.e. with differential rates of growth the LGCM-SR will be preferable. But it is noted that there are problems with recursivity because the change in the construct is now being estimated both in the within and between person models.

The RI-CLPM does not constrain the means to be equal over time. If this constraint were imposed, then the RI-CLPM would be nested in the LGCM-SR. Another model is the trait-state error (STARTS, for stable trait, autoregressive trait, and state) model. Kenny & Zautra (2001) decomposes variance into a time-invariant stable trait, an autroregressive trait which changes via an autoregressive process, and occasion specific state error (containing measurement error). Though this model is generally applied to univariate processes, it can easily be extended to multivariate processes. In the STARTS model, measurement error is accounted for whereas in the RI-CLPM it is not, thus STARTS is a generalized case of RI-CLPM wherein measurement error is accounted for. Though it does seem desirable to account for measurement error at given occasions, this can cause estimation problems if one does not have an adequate number of waves.

The Latent Change Score model (McArdle & Hamagami, 2001) is very similar to the LGCM-SR model with addition of latent variables for change scores, that are based on the difference scores accounting for measurement error. These change scores are modeled as a function of constant change by loading them on the slope term. They also contain proportional change, established by predicting the subsequent measurement occasion from the prior measurement occasion via the latent change score. In other words, the latent change scores result from creating an indirect path between measurement occasions through the latent change variable. The cross-lagged paths go from the measurement error corrected observed score at a prior time point predicting the change scores between that prior and subsequent measurement occasion. This model as can be seen is rather complex and the relationship between the CLPM and the latent change score has been explored in depth (Usami, Hayes, & McArdle, 2016; 2015).

A final model is the Latent State Trait model (LST; Steyer et al., 2015), in which an observed score is decomposed into measurement error, and a true score (containing a trait and state portion). The LST models requires multiple indicators at multiple occasions to estimate its respective components, but has been modified to handle single indicators (Luhman et al., 2011). This adaptation requires that we sacrifice the modeling of measurement error and the trait factor be modeled as a first order factor with loading being freely estimated. If we place the LST model into a bivariate model with cross-lags (as Luhmann et al., 2011 did) then the RI-CLPM is a special case of the LST, where the factor loadings are fixed to 1 across time.

As can be seen from above the RI-CLPM is one model in a battery of structural equation modeling approaches for controlling trait stability in addition to other models which will also account for trait change. If you recall from before there was some discussion of problems with estimating random effects models with lagged dependent variables. In the following I briefly demonstrate what a fixed effect analog of the RI-CLPM might look like.

2.5 LATENT GROWTH CURVE MODELS

In the prior models discussed we are primarily only dealing with trait like (time-invariant) stability, thus the control is on the intercept and change is mainly only being accounted for within the lags. This does not really account for growth in the more general sense, but rather establishes prediction across repeated measures, thus is more concerned with question of causality. These models are quite useful in this regard as fundamental assumptions of causal relations is that there be a temporal ordering such that the cause will always precede the effect and further that alternative causes can be ruled out. Though we can never fully rule out alternative hypothesized causal relations, accounting for autoregressive and cross-lag paths in a bivariate process bolsters the causal claims made about the causal relation between two processes. The claim is further bolstered when trait like stability is controlled out as well. However, these models aren't particularly informative in terms of gauging the actual trajectory of a developmental process. Developmental theories are generally interested in what growth actually looks like, and furthermore, variation in developmental trajectories amongst individuals is of central interest especially when exploring what factors may be influencing such differential growth between individuals. Additionally, within an individual we may see changes in growth over time that may be meaningfully attributable to the presence or absence of some factors in their lives.

To evaluate such variations between and within persons we must be able to represent some trajectory of development as well as capture variation between and within persons. These within person variations are time specific, in that such deviations are evaluated in reference to an average underlying growth trajectory. The desire to capture individual developmental trajectories motivates the interest in establishing suitable techniques for modeling and analyzing development in such a way that is consistent with our underlying theories and interest. This motivation has brought about a family of techniques known as growth curve analysis. A basic approach for this is to project change in a variable over time using regression models. We can then move into a multilevel framework to allow for variation in growth rates between persons, in the case of the random effect multilevel model we fit an average rate of change, and then allow for deviations in individual development with the inclusion of random rate of change. We can similarly capture such individual differences by allowing everyone to have their own fixed slope. Such approaches can be adapted to capture a whole range of trajectory shapes.

2.5.1 Univariate Linear Growth Curve Models

Latent variable modeling of growth (Bollen & Curran, 2006) proves to be a good and flexible approach to not just fitting curves but also integrating them with other modeling components available within structural equation modeling. For example, let's say we have multiple indicators for a construct. Within a latent variable approach to modeling growth it is possible to fit higher order growth curves from measurement models to capture the growth trajectory of such constructs. This capacity for building models in such modular fashion is one of the most desirable features of structural equation modeling. In the following we shall further discuss the latent growth curve approach and demonstrate how we can build upon growth curves to integrate cross-lagged panel models, in such a way that we can get a cleaner separating out of between person differences in developmental trajectories and within person processes of change, especially about the causal influence of two developmental processes on one another.

To begin we discuss the univariate growth curve null model. This model involves fitting a latent growth curve to a single indicator and including no exogenous predicting variables. This model is the fundamental building block for all latent growth curve models. The basic idea underlying the building of latent growth curves is to apply latent variables to repeated measures and set their loadings to capture the trajectory of change in a variable over time. To illustrate, let's say we have some variable y that varies across t time points for each individual i (y_{it}). As mentioned before we link these repeated measures to a trajectory via setting the loadings on latent variables in such a way as to reflect the change in the variable over time. For example, let's say we want to capture a linear growth trajectory, then for 6 equally spaced time points, we could capture this linear growth by establishing our loadings as λ=0, 1, 2, 3, 4, 5 for time points t=1,2,3,4,5,6 respectively and we capture the intercept by creating a latent variable to which we fix the loadings at all time points to be 1. The intercept is interpreted at the path loading on the slope factor that is equal to 0, so in the example above the intercept is interpreted at time point 1 (i.e., the initial occasion). The values placed on the slope path loadings will determine the intervals of times to which our slope is fit. In the above we are assuming that each point is equally spaced and that our intervals correspond to exactly the unit of time corresponding to our repeat measures.

If we wanted our slope rate to correspond to some factor of our unit of interval, then we could multiply the slope loadings by the inverse of the factor which we wish to scale our intervals too. For example, let's say we collected measures every other month, but we wish to interpret our slope in terms of months then we would set our slope loadings to λ=0,2,4,6,8,10,12. So far, we have considered our intervals as equally spaced, but this is not necessary. We can represent a slope directly in terms of the exact spacing of units if we wish. For example, often in education research we will have fall and spring measures, where perhaps the fall to spring measures are taken at equal interval in terms of months, but the spring to fall interval is not, so we might choose a set of loadings such as λ=0, 8, 12, 20, 24, 32. A range of loading series can be chosen from and the choice will be guided by the interpretation you wish to give to the intercept and slope.

We can define the intercept term as α and the slope term as β, we have the basic growth curve representation for the repeated measure as:

$$y_{it} = \alpha_{y_i} + \lambda_t \beta_{y_i} + \varepsilon_{y_{it}} \quad (5a)$$

where α_{y_i} represents an individual's intercept, and $\lambda_t \beta_{y_i}$ represents the growth trajectory of an individual with λ_t being the loadings at each t time point capturing the shape of the trajectory, and β_{y_i} capturing the individuals standing on the slope parameter. ε serves its usual role in representing the residual error for individual i at time t. The connection between the latent growth model and the random-effects growth model is apparent when considering the model for the terms for the individual's latent intercept and slope:

$$\alpha_{y_i} = \mu_{y_\alpha} + \delta_{y_{\alpha i}} \quad (5b)$$

$$\beta_{y_i} = \mu_{y_\beta} + \delta_{y_{\beta i}} \quad (5c)$$

The disturbance terms, δ, represent individual's deviances from the mean, μ, values for the intercept and slope. The compositional model has the representation:

$$y_{it} = (\mu_{y_\alpha} + \delta_{y_{\alpha i}}) + \lambda_t(\mu_{y_\beta} + \delta_{y_{\beta i}}) + \varepsilon_{y_{it}} \quad (5d)$$

The inclusion of the deviation terms makes it clear that we are looking at a random effect model, with a single covariate, namely the time variable. It is in these random components, i.e.

the disturbance terms, that we can capture between-person variability in intercept and growth. Expressing the fixed effects as the means for the intercept and slope, $E\begin{bmatrix}\alpha_{y_i}\\ \beta_{y_i}\end{bmatrix} = \begin{bmatrix}\mu_{y_\alpha}\\ \mu_{y_\beta}\end{bmatrix}$, and

having the covariance matrix for the random effects as $\psi = \begin{bmatrix}\sigma^2{}_{y_\alpha} & \\ \sigma_{y_\alpha y_\beta} & \sigma^2{}_{y_\beta}\end{bmatrix}$.

By fitting this covariance, we are saying that we believe that initial standing is related to growth, which is a typical expectation for a developmental process. The variance terms on the diagonal are the key values for capturing the between-person differences in growth, e.g. some people start out higher or lower at baseline and some show more or less change over time than others. The size of these terms will reflect how much people are differing, with higher values implying more between person difference. When it comes to residual errors we will also have some kind of structuring to consider. The generic specification will have independence amongst errors and in multilevel modeling is referred to as and residual covariance matrix which is represented as $var\begin{bmatrix}\varepsilon_{it}\\ \varepsilon_{it-s}\end{bmatrix} = \begin{bmatrix}\sigma^2_{\varepsilon_{it}} & \\ 0 & \sigma^2_{\varepsilon_{it-s}}\end{bmatrix}$ which implies that we only have diagonal elements

giving a residual variance for each time point and no correlation amongst these repeated measures. It may seem strange to assume that errors of repeated measures are uncorrelated, but, in the case of a growth curve it is not entirely infeasible to assume that the slope term is absorbing this dependence.

Within structural equation modeling there are different ways in which you can structure your residuals, with the unstructured residuals at the extremity where residuals at all time points are being freely correlated with one another and the most constrained model where the residual error for each time point is fixed as equal to the residual error at every other time point. The specification of equal variance across time points is a way of imposing homoscedasticity on a model which has favorable properties. If this assumption does not hold in the data, then model fit will be degraded and derived estimates become questionable. One of the key determinants in the tractability of different residual structure specifications is how far apart repeated measurement occasions are from one another. The further apart the less we anticipate such constricted assumptions to hold. Other factors will also influence this; for example, if an intervention occurred between repeated measures this may cause a change in residual errors. The benefit of more constrained residual structures is that it affords us degrees of freedom by requiring less parameters to be estimated. However, note again if the assumptions of given structures aren't tenable in the data then the quality of fitted model declines. This is the general structure of a latent growth curve, and we can call it the unconditional univariate growth curve with linear growth, unstructured covariance in the random effects, and an independent residual covariance structure. Most SEM software uses maximum likelihood estimation based on the multivariate normal distribution so it is implicit that the random effects and residuals are normally distributed. Figure 3 gives a visual representation of this model.

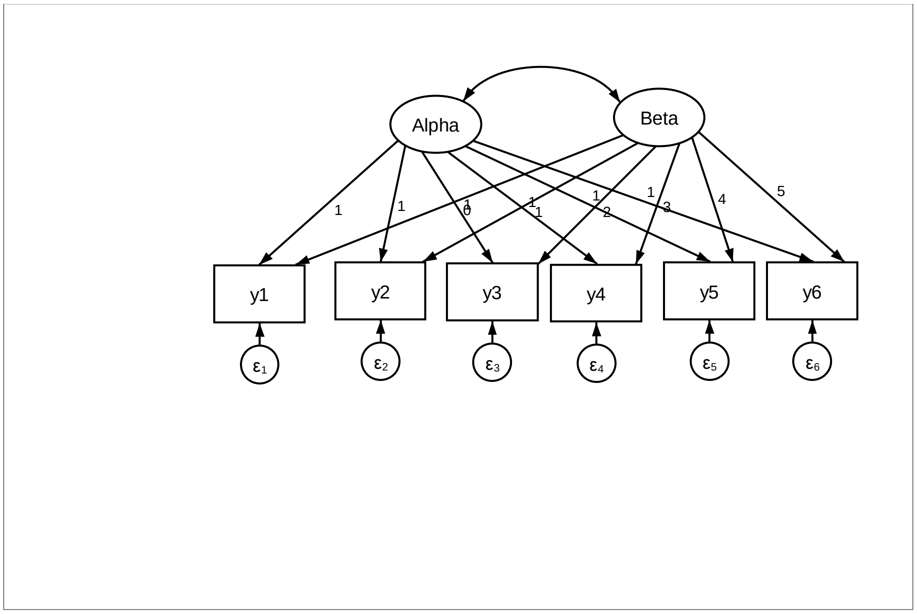


Figure 3. An unconditional latent growth curve model

2.5.2 Non-linear Growth Curves

Above what we have described is a linear trajectory. Through the manipulation of path loadings on the slope and inclusion of additional latent variables we can capture a whole range of functional (non-linear) forms for trajectories. The most general type of trajectory would be the unspecified growth curve, where we allow a trajectory to take a form that fits most closely to the data by leaving slope loadings to be freely estimated (Duncan et al., 1999). For identification purposes, we must fix two loadings. One path loading should be set to 0 to identify the intercept, the other path loading is relatively arbitrary, and sets a relative unit for the slope. The most common choices for this are, λ=0, λ_2, λ_3, λ_4, λ_5, 1, or, λ= 0, 1, λ_3, λ_4, λ_5, λ_6. Though this particular model is praiseworthy for its flexibility, the resulting trajectory of development may be hard to interpret. We sacrifice degrees of freedom in fitting this curve and so one needs to be mindful of how many degrees of freedom they have to spare in fitting this curve. This approach also runs the risk of becoming overfit, which is to say in approaching saturation we may have no variance left to explain. The concern with this is that we are following too closely to the noise in our data rendering such models fairly useless when it comes to fitting new data. Fundamentally, for our concerns, the problem with this is that we are establishing a theorized developmental process that is being too heavily influenced by a single sample. If one's concern is only to fit a curve to a single sample, then the unspecified growth is not problematic, but this is rarely the case in actual research.

A highly-related approach is to fit the slopes based on the mean differences, i.e. to fit the slope as running through the mean at every time point. Like the unspecified growth curve model this is just a matter of allowing slope loadings to be estimated. Either we can base this entirely by setting the slope shape by forcing the curve to fit through every time specific mean or estimate loading in terms of mean difference. The first approach will give slope loadings like such,

$\lambda = \overline{y_1}, \overline{y_2}, \overline{y_3}, \overline{y_4}, \overline{y_5}, \overline{y_6}$. Another approach using the mean is to fit the loadings for time 1 and time 2 as λ=0,1, then allow the remaining slope loadings to be estimated according to the mean difference between time 1 and time 2. This approach is highly similar to fitting the unspecified growth curve model where we chose to fix the time 1 and 2 loadings as λ=0, 1. We call this a modified mean spline model and it's resulting slope loading are as such

$$\lambda = 0, 1, \frac{\overline{y_3}-\overline{y_1}}{\overline{y_2}-\overline{y_1}}, \frac{\overline{y_4}-\overline{y_1}}{\overline{y_2}-\overline{y_1}}, \frac{\overline{y_5}-\overline{y_1}}{\overline{y_2}-\overline{y_1}}, \frac{\overline{y_6}-\overline{y_1}}{\overline{y_2}-\overline{y_1}}.$$

Another type of spline growth model is the piecewise spline. This involves fitting multiple latent linear slope variables that reflect different phases of development. For example, within the span of a study we may expect a high rate of growth earlier in the study, but during the later portion we expect this to be much more modest, so we might fit one trajectory for the earlier portion and another in the later portion. As non-linear modeling goes, piecewise splines have a nice interpretation in the context of development, in that we have broken a developmental process up into two different phases and have allowed growth rates to change within each specific phase. This is a commonly observed type of growth in developmental sciences. In this case, we now have multiple sets of slope loadings; for example, let's say we have two unique phases across our 6 time points, then we might have loading sets $\lambda 1=0,1,2,2,2,2$ for spline 1 corresponding to phase 1 and $\lambda 2=0,0,0,1,2,3$ for spline 2 corresponding to phase 2 of developmental.

The piecewise spline is similar to polynomial trajectories in that we model multiple latent slope variables to capture growth, but, they are very different in that piecewise splines capture phase specific growth trajectories using linear trajectories, whereas the polynomial curves utilize the latent variables to fit curvilinear trajectories in accordance with some n^{th} order polynomial up to the T-1 (where T=total time points) order. The most pronounced difference is that the polynomial curves don't necessarily fluctuate in direct accordance with phases of development, but rather lays out a more generalized trajectory of development. We generally do not want to go to high with polynomial orders, as each additional order of polynomial requires an added slope term as well as adding a layer of complexity in the interpretation of such a functional form. Quadratic and cubic curves remain relatively tractable and feasible for modeling developmental processes, i.e. polynomials of order 2 and 3. The general logic is that we have a latent variable for each order of polynomial, so for polynomial curve of order 2 (quadratic) we have a latent variable for linear growth with linear loadings, e.g. $\lambda=0,1,2,3,4,5$, and a latent variable for the 2^{nd} order polynomial (quadratic) with loading $\lambda^2 = 0, 1, 4, 9, 16, 25$. By extension if we had a cubic curve then we would add the cubic slope variable with loadings $\lambda^3=0, 1, 8, 27, 64, 125$. The components of latent growth modeling discussed thus far have centered on the univariate case, but we can expand our models to incorporate multiple growth processes together, each being permitted its own characteristics of growth curves.

2.5.3 Multivariate Growth Curve Models

The most basic of multivariate growth curve models is the bivariate parallel process model (McArdle, 1989) which essentially brings two growth curves together to capture concurrent and related processes of development. One could also postulate serial processes of development wherein one developmental process precedes the other in time. The model for this would most sensibly involve predicting one growth curve from the other. In this paper, we would develop from the parallel process model since this is most central to one of the overriding themes of this paper, namely, how do we determine causal dominance between two processes which we have measured across the same time span. For the bivariate parallel process growth, we will have two growth curve models:

$$x_{it} = (\mu_{x_\alpha} + \delta_{x_{\alpha i}}) + \lambda_t(\mu_{x_\beta} + \delta_{x_{\beta i}}) + \varepsilon_{x_{it}} \quad \text{(6a)}$$

$$y_{it} = (\mu_{y_\alpha} + \delta_{y_{\alpha i}}) + \lambda_t(\mu_{y_\beta} + \delta_{y_{\beta i}}) + \varepsilon_{y_{it}} \quad \text{(6b)}$$

All of the terms are defined as before for each respective process y and x. We relate these processes together through the unstructured covariance matrix:

$$var\begin{bmatrix}\delta_{x_{\alpha i}}\\ \delta_{x_{\beta i}}\\ \delta_{y_{\alpha i}}\\ \delta_{y_{\beta i}}\end{bmatrix}=\begin{bmatrix}\sigma^2_{\delta_{x_{\alpha i}}} & & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{x_{\beta i}}} & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\alpha i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\alpha i}}} & \sigma^2_{\delta_{y_{\alpha i}}} & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{y_{\beta i}}}\end{bmatrix}$$

Along the diagonal, we have the variances for each of the respective latent factors for slopes and intercepts while the off-diagonal captures the covariances amongst the slopes and intercepts. In this way, we are allowing individual's intercepts and slopes to be related, which is a good assumption to model given that we hypothesize some relation between these developmental processes. We now add some structuring to our residual covariance matrix to allow measures on the two constructs to correlate with one another within time points. Thus, we no longer have an independent residual covariance matrix but rather we represent it as such, now including two time variant variables:

$$var\begin{bmatrix}\varepsilon_{x_{t-s}}\\ \varepsilon_{x_t}\\ \varepsilon_{y_{t-s}}\\ \varepsilon_{y_t}\end{bmatrix}=\begin{bmatrix}\sigma^2_{\varepsilon_{x_{t-s}}} & & & \\ 0 & \sigma^2_{\varepsilon_{x_t}} & & \\ \sigma_{\varepsilon_{x_{t-s}}\varepsilon_{y_{t-s}}} & 0 & \sigma^2_{\varepsilon_{y_{t-s}}} & \\ 0 & \sigma_{\varepsilon_{x_t}\varepsilon_{y_t}} & 0 & \sigma^2_{\varepsilon_{y_t}}\end{bmatrix}$$

Along the diagonal, we find the residual error variances from time t-s to time t for each of our x and y variables, while within the off diagonal we see the residual errors are only correlating at specific shared time points. This is to say that the unexplained, remaining variance during measurement occasion between constructs has some relationship. In general, we will not correlate errors across time given that if measurements are separated in time this might imply some directedness in the relationship. In later models, we will begin exploring models that do more directly relate repeat measure to one another. Figure 4 gives a figurative representation of this model.

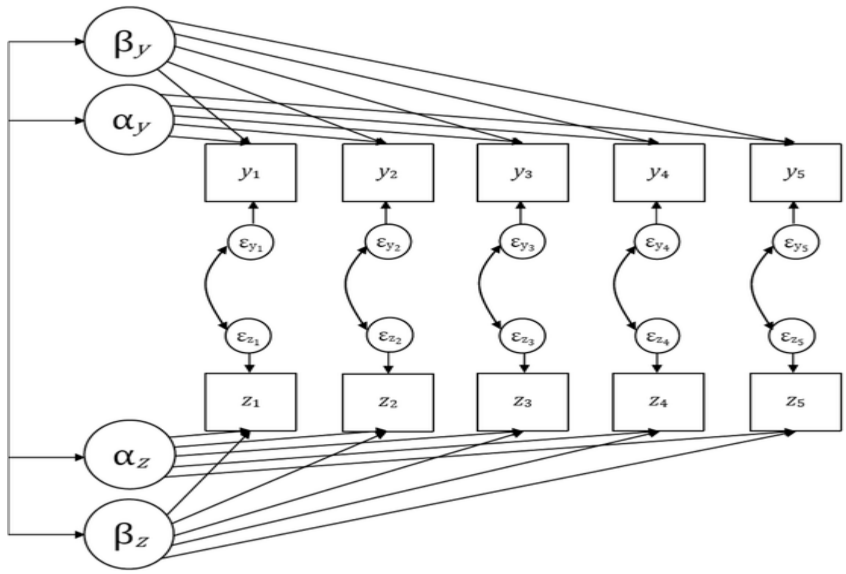


Figure 4. A parallel process bivariate growth curve model for 5 waves

2.5.4 Autoregressive Latent Trajectory

Since the univariate and bivariate growth curves as presented thus far are focused only on individual differences we aren't really gathering any information about how the time specific components of developmental processes. If we wanted to begin establishing relations between the constructs that would capture within person effects on development, we would need to start establishing relations amongst the time specific measures for individuals. It is at this point that we begin seeing the cross-lags and autoregression become integrated with growth curves. The autoregressive latent trajectory model (Curran & Bollen, 2001; Bollen & Curran, 2004) essentially fits a growth curve on a cross lagged panel model. Additionally, similarly as is done with fixed effects models, we allow the initial measurement of our variables to correlate with the person -specific intercepts and slopes. The resulting covariance matrix is given as:

$$var\begin{bmatrix}\delta_{x_{\alpha i}}\\ \delta_{x_{\beta i}}\\ \delta_{y_{\alpha i}}\\ \delta_{y_{\beta i}}\\ x_{i1}\\ y_{i1}\end{bmatrix}=\begin{bmatrix}\sigma^2_{\delta_{x_{\alpha i}}} & & & & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{x_{\beta i}}} & & & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\alpha i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\alpha i}}} & \sigma^2_{\delta_{y_{\alpha i}}} & & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{y_{\beta i}}} & & \\ \sigma_{\delta_{x_{\alpha i}}x_{i1}} & \sigma_{\delta_{x_{\beta i}}x_{i1}} & \sigma_{\delta_{y_{\alpha i}}x_{i1}} & \sigma_{\delta_{y_{\beta i}}x_{i1}} & \sigma^2_{x_{i1}} & \\ \sigma_{\delta_{x_{\alpha i}}y_{i1}} & \sigma_{\delta_{x_{\beta i}}y_{i1}} & \sigma_{\delta_{y_{\alpha i}}y_{i1}} & \sigma_{\delta_{y_{\beta i}}y_{i1}} & \sigma_{x_{i1}y_{i1}} & \sigma^2_{y_{i1}}\end{bmatrix}$$

The residual covariance matrix is as before with residuals correlated within time across constructs, except now we no longer have residual correlation at time 1 (i.e., t-s) because this is now exogenous and is presented in the covariance matrix with the latent intercepts and slopes. The score decompositions can be given as:

$$x_{it} = (\mu_{x_\alpha} + \delta_{x_{\alpha i}}) + \lambda_t\left(\mu_{x_\beta} + \delta_{x_{\beta i}}\right) + \rho_{xx}x_{it-1} + \rho_{xy}y_{it-1} + \varepsilon_{x_{it}} \quad (7a)$$

$$y_{it} = (\mu_{y_\alpha} + \delta_{y_{\alpha i}}) + \lambda_t(\mu_{y_\beta} + \delta_{y_{\beta i}}) + \rho_{yx}x_{it-1} + \rho_{yy}y_{it-1} + \varepsilon_{y_{it}}$$

(7b)

Figure 5 gives a representation of this model. Though this model gives us some information about more time specific effects in the cross-lagged and autoregressive parameters, it does not clearly separate out the between and within-person aspects of the developmental processes. One may note when looking at the figure that it appears that the effect of one construct on the other is being mediated through the relationships amongst the individual indicators. This sort of mediation dynamic is still being maintained at the level of individual differences such that the within and between person levels are still conflated with one another and the effects being captured within the cross-lag model is that earlier measures on one construct are causally linked to later measures on the other construct.

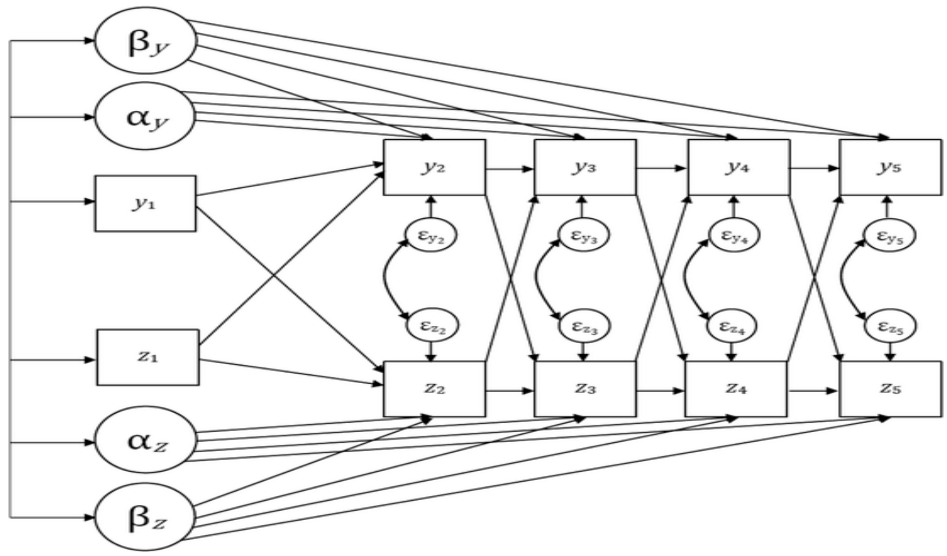


Figure 5. An Autroregressive Latent Trajectory Model with Cross-Lagged effects

2.5.5 Latent Growth Curve Model with Structured Residuals

When we maintain in theory that our developmental processes have a unique between and within person component we will need a model that makes a clearer separation of effects occurring at the between and within person level. The autoregressive latent trajectory model is limited in this regard. Conceptually speaking, it is easy to see that the best indicator of the within person component can be found in the residuals, which conceptually represent that which remains unexplained in an observed variable once we account for individual differences, i.e. the residual represents time t specific deviations for an individual i on a given process (x or y). Thus, we turn to the idea of structuring residuals to cleanly pull out the within person level of effect present in a development process (Curran et al., 2014). By structuring the residuals, we are now able to thoroughly disaggregate the between and within person effects, where the between person level is being captured in the latent growth curve portion and the within person level is being captured in the structured residuals. Particularly by structuring the residuals into a cross-lagged panel model we will have the elements required to assess temporal stability in a construct and the causal dominance relationship between two constructs at the within-person level.

This logic hearkens us back to the RI-CLPM as presented in a prior section, the difference being that now we are also accounting for between person variability in development. Because we are fitting the model to the residuals we will not be influencing the fixed effects (i.e. the mean structure) in the growth curves as we did with the autoregressive latent trajectory model which placed the cross lagged structure on the observed variables. Because our modeling technique involves fitting growth curves then structuring the residuals the model can be simply referred to as the Latent Growth Curve with Structured Residuals (LGCM-SR). When considering that the addition of the structured residuals is simply an extension on the LGCM. The LGCM-SR is a generalized model in which both the multivariate and univariate growth curves are nested. This is a nice feature of this modeling approach as it allows us to employ likelihood ratio tests to evaluate how much improvement we get by separating out between from within person levels of analysis through the fitting of the extra parameters.

Recall that, for example, $y_{it} = \alpha_{y_i} + \lambda_t \beta_{y_i} + \varepsilon_{y_{it}}$ where $(\alpha_{y_i} + \lambda_t \beta_{y_i})$ represents the beween person aspect of the observed variable y_{it} then through a simple reworking we see that by removing the between person component from the measure we are left with

$\varepsilon_{y_{it}} = y_{it} - (\alpha_{y_i} + \lambda_t \beta_{y_i})$ which highlights that the residual is capturing the within person component as the individual's time specific deviation from the growth curve (between person) component. From here we can impose meaningful structure to the residuals that will prove informative as concerns the within-person process. As we have explored before one thing to structure out is the temporal stability in repeated measures, which is accomplished by choosing to autoregress the residuals on one another across time instead of putting them into some residual covariance matrix where they will either be treated as independent or structured in some meaningful ways.

Regardless of the residual matrix structure the fact remains that we don't propose a directedness in effect, but such an assumption is not entirely accurate regarding developmental processes where we anticipate that prior measures serve as strong predictors of later predictors. Hence it makes sense to impose a directed relation wherein prior values on a construct predict subsequent measures. Note that we are not gauging pure stability in a construct through the auto-regressed residuals, as some of this stability will be accounted for in the latent intercept factor, similar to the intercept factor in the RI-CLPM referred to as being a trait like stability. Additionally, we are also absorbing some of the carry-over effect from one prior measure to the next into the slope parameter. Thus, the best way of considering the autoregressed residuals in this model are as the carry over effects from one occasion to the next net the between person differences in individual's growth trajectories. We now can represent the residuals as

$$\varepsilon_{y_{it}} = \rho_{yy}\varepsilon_{y_{it-1}} + v_{y_{it}} \quad (8a)$$

$$\varepsilon_{x_{it}} = \rho_{xx}\varepsilon_{x_{it-1}} + v_{x_{it}} \quad (8b)$$

where the *v* term is now taking the former role of the residual, so we will refer to this term as the residual error. These definitions of the residual will remind us of the RI-CLPM where we did this same thing. With this as a foundation we can been entering our cross-lag pathways to have a way to simultaneously estimate both the between person differences in development processes over time (parallel process growth) along with the within person inter-lag relations between processes. Figure 6 gives a depiction of this model.

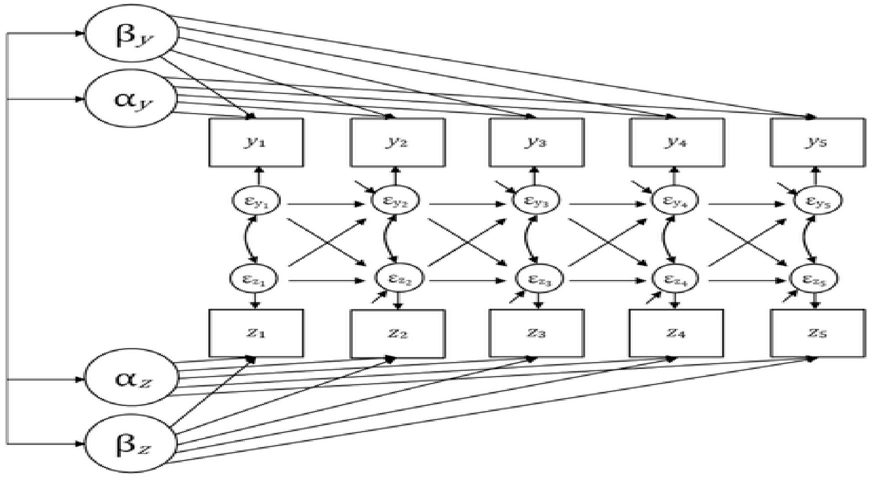


Figure 6. Latent Growth Curve Model with (ARCL) Structured Residuals

Once we enter the cross-lag pathways and integrate the components together. We represent the LGCM-SR as:

$$x_{it} = (\mu_{x_\alpha} + \delta_{x_{\alpha i}}) + \lambda_t \left(\mu_{x_\beta} + \delta_{x_{\beta i}}\right) + (\rho_{xx}\varepsilon_{x_{it-1}} + \rho_{xy}\varepsilon_{y_{it-1}} + v_{x_{it}}) \tag{8c}$$

$$y_{it} = (\mu_{y_\alpha} + \delta_{y_{\alpha i}}) + \lambda_t \left(\mu_{y_\beta} + \delta_{y_{\beta i}}\right) + (\rho_{yy}\varepsilon_{y_{it-1}} + \rho_{yx}\varepsilon_{x_{it-1}} + v_{y_{it}}) \tag{8d}$$

with between person covariance matrix as before with the standard bivariate LGCM:

$$var \begin{bmatrix} \delta_{x_{\alpha i}} \\ \delta_{x_{\beta i}} \\ \delta_{y_{\alpha i}} \\ \delta_{y_{\beta i}} \end{bmatrix} = \begin{bmatrix} \sigma^2_{\delta_{x_{\alpha i}}} & & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{x_{\beta i}}} & & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\alpha i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\alpha i}}} & \sigma^2_{\delta_{y_{\alpha i}}} & \\ \sigma_{\delta_{x_{\alpha i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\beta i}}\delta_{y_{\beta i}}} & \sigma_{\delta_{x_{\alpha i}}\delta_{x_{\beta i}}} & \sigma^2_{\delta_{y_{\beta i}}} \end{bmatrix}$$

with the residual error covariance matrix being expressed similarly to the residual matrix before with correlation between constructs being restricted as time specific:

$$var\begin{bmatrix} v_{x_{t-s}} \\ v_{x_t} \\ v_{y_{t-s}} \\ v_{y_t} \end{bmatrix} = \begin{bmatrix} \sigma^2_{v_{x_{t-s}}} & & & \\ 0 & \sigma^2_{v_{x_t}} & & \\ \sigma_{v_{x_{t-s}}v_{y_{t-s}}} & 0 & \sigma^2_{v_{y_{t-s}}} & \\ 0 & \sigma_{v_{x_t}v_{y_t}} & 0 & \sigma^2_{v_{y_t}} \end{bmatrix}$$

All of the descriptions for the elements of these simply involve a recasting of model components into the framework where we structure residuals into a cross-lag panel model instead of the observed variables. As is clear at this point, in doing this, we account for the within-person level via the structured residuals and the between-person level remains in the growth curves.

At this point it is important that we relate the RI-CLPM to the LGCM-SR. As given before the RI-CLPM is formulated as:

$$x_{it} = \mu_{x_t} + \alpha_{x_i} + \rho^*{}_{xx}\varepsilon^*{}_{x_{it-1}} + \rho^*{}_{xy}\varepsilon^*{}_{y_{it-1}} + v^*{}_{x_{it}} \qquad (2e)$$

$$y_{it} = \mu_{y_t} + \alpha_{y_i} + \rho^*{}_{xx}\varepsilon^*{}_{y_{it-1}} + \rho^*{}_{yx}\varepsilon^*{}_{x_{it-1}} + v^*{}_{y_{it}} \qquad (2f)$$

with residual covariance matrix:

$$\psi = \begin{bmatrix} \sigma^2_{\varepsilon^*_{x_{t-s}}} & & & & & \\ \sigma_{\varepsilon^*_{x_{t-s}},\varepsilon^*_{y_{t-s}}} & \sigma^2_{\varepsilon^*_{y_{t-s}}} & & & & \\ 0 & 0 & \sigma^2_{v^*{}_{x_t}} & & & \\ 0 & 0 & \sigma_{v^*{}_{x_t},v^*{}_{y_t}} & \sigma^2_{v^*{}_{y_t}} & & \\ 0 & 0 & & & \sigma^2_{\alpha_x} & \\ 0 & 0 & 0 & 0 & \sigma_{\alpha_x,\alpha_y} & \sigma^2_{\alpha_y} \end{bmatrix}$$

We note that in the covariance matrix, where we have included the intercepts as well as the residual errors, there is no real difference between the two models aside from the fact that with the addition of the slope terms we have additional covariances between slopes and between intercepts and slopes. We allow the between person components to freely covary. Beyond the obvious difference created by the addition of a random slope components, we also find that the intercept component in the RI-CLPM only contains one term. This term is the variance of the intercept and is analogous to the random intercept variance from the LGCM-SR. Unless means are treated as constant across time, we do not have a mean intercept in the RI-CLPM as we do in the LGCM-SR; instead we have time specific means μ_t. It is by fitting these time specific means that the RI-CLPM accounts for drift in the measurements across occasions, while this role is serviced by the slope mean in the LGCM-SR, especially in the case where we fit our curves as mean spline or unspecified. Since we are centering our individual values around the time specific means in the RI-CLPM, we also have a case where the intercept mean value is zeroed out as every time point's value for the mean is zero relative to the individual deviations representing our within person component.

In saying the mean relative to individual scores is zero across the within person component, we are saying the between person mean across time is set at zero, so the models are largely doing the same things by changing the place where the means are estimated. To reiterate the RI-CLPM is incorporating the means into the within person, time specific component whereas the LGCM-SR is incorporating the means into the between person component. The RI-CLPM is therefore not nested under the LGCM-SR properly speaking, but the two models are essentially accomplishing the same thing, though it is fair to say that the LGCM-SR is pulling more variance into the between person component through the addition of the slope parameters. This applies not only to the univariate components but also to the cross-construct component where the covariance of the slope parameters across constructs can be expected to absorb more of the cross-lagged components given that the cross-lags are capturing across time relations between the constructs just as the slope parameters are.

3.0 METHODS

Chapter 3 begins with a demonstration of the aforementioned models to illustrate how the different models can lead to different and sometimes divergent results. Following this empirical example, we will present a simulation study designed to evaluate the relative performance of the models under various conditions pertaining to characteristics of change over time and the nature of the reciprocated relationship. The efficacy of the approach to the simulation study will be demonstrated in the data generation and validation section.

3.1 EMPIRICAL EXAMPLE WITH LONGITUDINAL STUDY OF AMERICAN YOUTH

To demonstrate and compare amongst various cross-lagged panel models that do and do not disaggregate between and within-person effects we fit these models to the bivariate processes of task value and self-concept as pertains to mathematics for 7^{th} through 12^{th} graders. The primary interest is in evaluating model fit and the causal dominance amongst the processes that each model implies. The constructs of math task value and math self-concept are central to motivational research in mathematics achievement (Wigfield & Eccles, 2000). From expectancy-value theory the idea of subjective task value in math pertains to one's beliefs about the benefits and utility of learning mathematics and self-concept pertains to one's beliefs about how good they are in mathematics. In terms of causal precedence, the question is whether belief in one's ability in math leads to a greater sense of value in mastery over mathematics, or vice versa, or if there is any kind of causal precedence at all.

3.1.1 Data

The data utilized in this study are derived from the 2007 to 2011 cohort of the Longitudinal Study of American Youth (LSAY) which is a national sample of 3,116 seventh graders who were enrolled in public school. For the following analysis, only complete cases were used, resulting with a sample with N=1,060. This study evaluated student's attitudes towards mathematic achievement each year from 7^{th} to 12^{th} grade. The task value variable can range from 0 to 25, while self-concept could range from 0 to 15. Subjective task value was based on student's responses on Likert-type scales (1=Strongly Disagree to 5= Strongly Agree) to survey items such as (e.g., "I enjoy math.", "Math is useful in everyday life.", etc.) (Eccles et al., 1997). Similarly, student's self-concept in their mathematic ability was based on student's responses on Likert-types scales (1=Strongly Disagree to 5=Strongly Agree) to survey items such as (e.g., "I am good at Math.", "I usually understand Math.", etc.) (Bleeker & Jacobs, 2004). On both scales, higher scores reflect greater task value and self-concept respectively. The math ability self-concept had good reliability with α=0.80, and the task value scale had relatively good reliability with α=0.75.

3.1.2 Analytic Plan

All models were fit in MPlus 7.4 (Muthen & Muthen, 1998-2015). In all, seven different models examined the bivariate relations amongst task value and self-concept. Models were constrained to have autoregressive parameters within a construct equal across lags as well as holding each task value->self-concept and self-concept->task value cross-lag parameters equal across lags. Additionally, exogenous components could freely covary, (e.g. residuals at the first-time point, intercepts and slopes). For the reciprocated models, endogenous time specific residuals across constructs had fixed covariance at all time points. At first, we fit a cross-lag panel model which does not disentangle within and between person effects but rather only considers rank ordering of individuals over time. Second, we fit a CLPM with time specific means estimated, which is the same as doing a cross-lag panel model on time-specific mean centered variables (i.e., essentially a group mean centering where occasion is the grouping variable). The second model is assessing the relationship amongst the individual's time specific deviations at given lags. Hence Models 1 and 2 are operating similarly to one another, namely, evaluating relative standing of individuals at given time points. In Model 3, we fit a CLPM directly to repeated measures that are centered around their individual specific means across time, i.e., person-mean centered variables. In this model, we aren't really disentangling between and within person effects. Instead, we are just examining the variables over time in relation to an individual's mean score across time. Hence, we are only evaluating relative standing of an individual in relation to their overall standing.

Extending from the CLPM with estimated means, a random intercept is fit so that trait like stability can be separated from within person changes across time. In Model 4, RI-CLPM, the between and within person effects are being disentangled. In Models 5 & 6, we evaluate the latent growth curve model with structured residuals (LGCM-SR), which not only disentangles trait-like stability between individuals but also examines inter-individual differences in change over time by fitting a random slope in addition to the random intercept. Residuals are then structured in such a way as to allow the autoregressive and cross lag effects of within person change amongst task value and self-concept to be estimated. This same strategy is used in both RI-CLPM and LGCM-SR. In fitting the LGCM-SR we first consider each latent growth process separately. To gather a sense of the most adequate shape of the trajectory we first fit the linear trajectory (Model 5) and then the unspecified trajectory (Model 6), which allows the curve to take any form. The two trajectories are then brought together into two different LGCM-SR models: one with a linear trajectory and the other with an unspecified growth curve.

Model 7 is a fixed intercept model (FI-CLPM) which disentangles between and within person processes as well as looking at autoregressive, cross-lag (ARCL) series. However, this model will not be truly bivariate in the sense that the other models are, but rather looks at each ARCL separately. For example, first we fit the model to evaluate the prediction of task value from self-concept (i.e., task value is the lagged dependent variable, while self-concept is the time-varying, cross-lag predictor) and then fit the model predicting self-concept from task value. Some of the unique features of this model involve correlating the first measure of the lagged dependent variable series with the latent intercept, correlating the intercept with the values of the time-varying, cross-lag predictor, and allowing the time-varying predictor measures to correlate with one another as well as the first measure in the lagged dependent variable series. The lagged dependent variable at all time points after the first is independent of the predictor terms, including the cross-lag and intercept. Placing the intercept as correlated with the predicting variables establishes it as a fixed effect independent of the lagged dependent variable.

Upon fitting these models, we evaluate them in terms of model fit to gauge which model appears to be doing the best at fitting the data. We also examined the cross-lag parameters to see (1) which process is indicated as being causally dominant and (2) what is the size of these effects in relation to one another. Of key interest in the between person components is the size of the variance which implies the extent to which individuals are different from one another in trait-like stability or inter-individual differences in change. Finally, where relevant, we examined the covariance in the between person components.

3.2 METHODS FOR SIMULATION STUDY

In the following simulation studies, we will be comparing the RI-CLPM with the LGCM-SR to evaluate performance of the models under various conditions. Due to the nature of its current specification, the FI-CLPM approach is not included in these simulation studies because it does not align well enough to the RI-CLPM and LGCM-SR to make comparisons particularly clear. Additionally, the CLPM specifications are also not included as prior studies (Hamaker et al., 2015; Berry & Willoughby, 2016) have made these comparisons.

Though results are not reported in the applied example above, the models were also fit to ECLS-K data for math and reading scores from Kindergarten to 8^{th} grade. Because time points were not equally spaced and the trajectory was extremely non-linear, this was not used as the illustrative example. However, the estimates derived from it were found to be more favorable for basing the parameter values on. Much of this pertains with the fact that more variance is observed in the ECLS-K data. When using LSAY data for parameters it was found that altering variance and covariance often led to negative variance estimates or correlation amongst the latent variables greater than one. Moreover, the LSAY variables exhibit relatively small change over time, while the ECLS-K variables exhibit quite a bit more change over time. This latter attribute, in combination with the distinctly non-linear trajectory of the ECLS-K variables, allows for a better evaluation of models in relation to questions concerning the functional form of growth.

3.2.1 Study Design

Across conditions, results from three models will be compared: the RI-CLPM, the LGCM-SR with a linear trajectory, and the LGCM-SR with an unspecified trajectory. Data will be generated from each respective model and subsequently analyzed by each of the models, including the model that corresponds to the generating model. Though we anticipate that the generating model will be fit best by its corresponding analytic model, the reason for fitting this model is to give a sense of the performance of the other models relative to the accurately specified model.

The central interests of this study pertain to how model fit and parameter bias are affected by the shape and variability of change over time, the conclusions we make about causal dominance, and the relationship between processes over time. Consequently, the independent factors will pertain to the functional form of the slope, the slope variances and covariance, and whether a dominant process is present or not.

To have variability in growth over time that reflects typical data dealt with in developmental and educational research, the slope variance conditions are based on slope variance estimates on trajectories in Math and English over the course of elementary school. The baseline slope variance estimates are derived from the ECLS-K, and we will be used to represent a middle level for variance. To reflect a lower variance condition, we can reduce the baseline variances by half of the baseline. To reflect a higher variance, we can double the baseline variance. These values will allow us to gauge changes as we move from lower to higher variance over a satisfactorily wide range. The slope covariance will be established based on Cohen's criteria for small, medium, and large effect sizes in terms of correlation values. This is used by convention but will still allow a considerable range of covariance magnitudes to evaluate performance of models across conditions. The trajectories are guided by the generating models: the linear LGCM-SR is a basic linear trajectory as would typically be fit to data, while the unspecified LGCM-SR slope trajectory and the time specific means for the RI-CLPM are also based on K-8th grade Math and English scores from the ECLS-K data. The final condition to be altered pertains to the causal dominance in the cross-lags. For this we have two basic conditions one where we have no dominance and one where dominance is present.

This results in 2 specifications of the RI-CLPM, dominance vs. no dominance, being analyzed by the three models. For the LGCM-SR we will have 2 sets of slope loadings X 3 slope variance conditions X 3 slope covariance conditions X 2 dominance conditions for a total of 36 specifications to be analyzed by the 3 models. This gives us (2 X 3) + (2 X 3 X 3 X 2 X 3)=114 sets of simulation conditions, and each condition has 1,000 replications. The different datasets generated from this design will be evaluated in terms of (1) the extent to which admissible and convergent solutions are derived; (2) model fit criteria; and, (3) estimation of the cross-lag parameters. Analyses of Variance will be used to understand the influence of the various factors on the various outcomes.

3.2.2 Fixed Factors

We will utilize six time points, which is both the number of time points presented in Curran et al. (2014) demonstration of the LGCM-SR and is also the number of time points given in the ECLS-

K data set. The sample will be set at N=5,000, to reflect a large enough sample to acquire trustworthy estimation given the number of parameters contained in these models. This is about a quarter of the size of the ECLS-K without considering missing values, but because of attrition we generally anticipate a loss in subjects, and over such a long span of time we may expect this to be severe.

Other fixed values are derived from analyses on the ECLS-K, with the specific numeric values identified in the *Data Generation and Validation* sections. As mentioned above, given that both models fit an intercept, the variance will not be of great interest in this study, and thus intercept variances and covariance will be fixed. Correlation effect size guidelines based on Cohen's criteria will be used to determine covariance values. More details on this procedure are given in the Data Generation and Validation sections. The covariance in intercepts will be fixed according to a correlation value of 0.3, reflecting a moderately strong correlation between intercepts. Similarly, we will fix the covariance of slope to intercept within a construct based on a correlation value of 0.3 to reflect a moderate relationship. Since we would expect the covariance of slope to intercept across constructs to be of smaller magnitude we will use a correlation value of 0.1 to set the covariance.

Residual variances are based on values derived from fitting a model allowing the time 1 variances to be freely estimated with the remaining t=2 through 6 variances constrained to be equal within a construct. As was done for the intercept covariance, the time specific residual covariances are fixed based on a correlation value of 0.3 reflecting a moderate relationship. To establish analogous stability within constructs across time that indicate upward prediction from one time point to the next, we use an autoregressive path loading of 1.2. The choice of having a

20% increase in value across lags relates to having 5 lags and hence a 100% added value over the 5 lags.

3.2.3 Independent Factors

As mentioned above we will have 2 specifications of the RI-CLPM, one reflecting no dominance and one reflecting causal dominance. As it is not found to be of great interest which process is dominating which since no substantive interpretation is being applied to the variables, in the dominance condition, we will have Y dominate X by a factor of 4. The dominant cross-lag loading will have a value of 0.8 and the dominated cross-lag loading will be 0.2. The value of 0.8 is chosen to reflect a situation where a unit change in the prior value of X corresponds to nearly an entire unit increase in the subsequent value of Y. In the non-dominance condition, we will reflect a situation where both constructs have an equally strong prediction of one another at subsequent lags, hence both the X to Y and Y to X cross-lag parameters will have a value of 0.8.

The specifications for the LGCM-SR will have 2 sets of slope loadings, one indicating linear growth (λ=0, 1, 2, 3, 4, 5) as is conventionally done, and the other reflecting non-linear growth. The loadings for the non-linear trajectories were acquired by fitting an unspecified growth curve to the ECLS-K data, we set the time 1 loading at 0 and the time 2 loading at 1, which must be done for identification purposes, and then to map onto the linear the time 6 loading was bounded at 5 and the preceding loadings were adjusted accordingly. For the X variable, this results in (λ= 0, 1, 1.175, 2.35, 4.7, 5) and for the Y variable (λ=0, 1, 1.09, 2.875, 4.37, 5).

In terms of the slope variance, as discussed above, we have 3 conditions ranging from low to high. The low condition is based on reducing the baseline variance by half, the medium condition is the baseline variance, and the high variance condition doubles the baseline variance. This is done to give a sufficient range for evaluating the behavior of the models as we move from lower to higher variance. In numbers, we have for the X variable, low slope variance=4, medium slope variance=8, and high slope variance=16; and for the Y variable, low slope variance=5, medium slope variance=10, and high slope variance=20. The slope covariance will have 3 conditions as well, reflecting a range for lower to higher slope covariance. These values are calculated by applying a correlation matrix to the vector of variances derived from the ECLS-K (see details in the Data Generation and Validation sections).

In establishing the size of the covariance, we use Cohen's criteria for correlation effect sizes, thus we have a low correlation = 0.1, a medium correlation=0.3, and a high correlation=0.5. The cross-lag conditions for the LGCM-SR generating model will be the same as with the RI-CLPM generating model.

3.2.4 Dependent Variables

The different models given above will be compared in terms of various model fit statistics. Three general classes of model fit indices will be considered: information criteria, absolute fit criteria, and relative fit criteria. Namely, we want to know how often the model fit indices select the correct model.

Two of the most commonly used information criteria are Akaike Information Criteria (AIC; Akaike, 1973) and Bayesian Information Criteria (BIC; Schwarz, 1978). Information criteria favor overall goodness of fit in a model in terms of the likelihood function while penalizing for model complexity due to the number of parameters estimated. In this way, information criteria aid a researcher in determining more parsimonious models. The information criteria given above are closely related, but vary in terms of the penalty they give for the number of parameters fit. The AIC can simply be phrased as multiply the parameters by then subtracting negative twice the log likelihood, AIC=2k-2ln(L), thus increases in the likelihood function (L) are offset by increases in parameters(k). If adding more parameters doesn't substantially improve the likelihood function, then only small reductions or increases in information criteria will ensue. From this we can see that with information criteria we want smaller values for a model relative to another. BIC functions under the same logic but applies a more stringent penalization than the AIC. Recall that with AIC we simply doubled the parameters estimated to penalize the likelihood function, thus only the number of parameters was accounted for in the penalty. BIC incorporates the sample size into this penalty, because sample size (n) is an important factor in setting the likelihood function. The formula now becomes BIC=ln(n)*k-2ln(L).

Common indices suggested for use with Latent Growth Models (Curran, Obeidat, and Losardo, 2010) are the comparative fit index (CFI; Bentler, 1990), the root mean square of approximation (RMSEA; Steiger & Lind, 1980), and the standardized root mean square residual (SRMR). These model fit indices can be classified into two main varieties, absolute fit of a model and relative fit of a model. RMSEA and SRMR are examples of absolute fit while CFI and Tucker-Lewis Index (TLI) are based on the fit of a model relative to the null model.

The SRMR pertains directly to the covariance matrices representing the model. Specifically, if we took the difference between the proposed models covariance matrix and that observed in the data we would have a matrix of residuals, SRMR represents this difference in standardized values, i.e. SRMR is based on the difference between the observed and predicted correlation matrices. Because of this, we know that an SRMR value of 0 is a perfect fit. The RMSEA also penalizes the likelihood χ^2 in relation to parameters and sample size. The formula is RMSEA=$\sqrt{}(\chi^2$ – df)/ $\sqrt{}$(df(N-1)), such stats assume that a fit of zero is the best fit and so are evaluating “badness” of fit in terms of how far the value is from 0. Naturally, the smaller these values are the better the fit is assumed. Cut values are generally controversial but MacCallum, Browne, and Sugawars (1996) propose that the values of .01, .05, and .08 be used to indicate excellent, good, and mediocre fit respectively.

CFI compares between the model set forth and the null model and determines fit based on this. The strategy here is to penalize the likelihood χ^2 for each additional parameter estimated by this formula Null(χ^2 -df) – Proposed(χ^2-df)/ Null(χ^2 -df). Interpretively the closer this value is to one the better we assume model fit to be, note here also that we set the value to 1 if it goes over 1 and 0 if it goes under 0. The TLI is very similar to the CFI, except it has a more conservative penalty based on the ratio of the likelihood χ^2 reduction to the added parameters: (Null(χ^2/df)-Proposed(χ^2/df)) / (Null(χ^2/df) – 1).

Another criterion for comparing amongst the models evaluated in this study is to consider biases in estimation. Since the main purpose in employing these models is determining causal dominance amongst multivariate processes, the main parameter biases of interest pertain to the cross lagged parameters. When considering bias in cross-lag estimates the key issues concern the bias in the point estimate for the cross-lag path, and the bias in the standard errors of estimating the cross-lag path.

The average relative bias in cross-lag path estimation across replications is a general indicator for the overall severity in bias. This criterion is defined as: $Mean\ Relative\ Bias = \left({}^{(\hat{\Theta} - \Theta)}/_{\Theta} \right) /_{R}$, where the numerator represents the difference between the estimated and true parameter divided by the true parameter value and in the denominator, is the number of replications. This criterion gives us two important pieces of information: (1) the magnitude of estimation bias, and (2) the direction of this bias, i.e. over vs. under estimation. When considering the bias in the standard errors of the cross-lag path estimate we can take the Monte Carlo standard deviation (σ_{Θ}^{MC}) of the cross-lag coefficient estimate as our true sampling variance for the estimate. With this we can calculate relative bias in the standard error as $\left(\frac{\left((\widehat{s_{\Theta}} - \sigma_{\Theta}^{MC})/\sigma_{\Theta}^{MC}\right)}{R}\right)$. The bias in the standard errors is important when considering the adequacy with which our estimates are reliable.

Another aspect of point estimate bias we are interested in pertains to how well the estimation represents the true nature of the causal dominance relationship. For this we can evaluate an index representing the causal dominance relationship. To measure the averaged factor of causal dominance across replications we can use the ratio $\frac{\left|\widehat{\theta_{xy}}/\widehat{\theta_{yx}}\right|}{R}$ when the value in the numerator is greater than one then we conclude that Y dominates and when less than one we conclude that X dominates and when equal to one we determine no dominance. In this study, we have 2 causal conditions, (1) X dominates Y by a factor of 4, and (2) no dominance, i.e. ratio equals 1. This index is important in considering the extent to which we will derive correct conclusions about the relationship between two processes.

3.3 DATA GENERATION AND ANALYSIS

Data Generation: Three data generation models were used a Random Intercept Cross Lagged Panel Model (RI-CLPM; Hamaker et al., 2015), Latent Growth Curve Models with Structured Residuals (LGCM-SR; Curran et al., 2014), one with a linear trajectory and another with an "unspecified" trajectory. The general steps for generating data sets follows as such: First, specify an error correlation matrix using PROC IML in SAS 9.4. Determining the magnitude of correlations in accordance with Cohen's (1988) effect size criteria with r=0.1 corresponding to a small relationship implying 1% shared variance, r=0.3 corresponding to a moderate relationship implying 9% shared variance, and r=0.5 corresponding to a large relationship implying 25% shared variance. Second, the respective models are fit in MPlus 7 (Muthen & Muthen,1998-2015) to math and reading IRT scores from Kindergarten to 8^{th} grade contained in the ECLS-K (Early Childhood Longitudinal Study) sample to acquire mean and variance values (and in the case of the 'unspecified' trajectory, slope loadings) reflective of real-world educational data. Third, the variance values are fed into PROC IML SAS 9.4 to calculate a covariance matrix from the error correlation matrix specified in step 1. This covariance matrix in conjunction with the mean values is then used to generate a sample of N=1,000 observations derived from a multivariate normal distribution. The variables generated from this prior step represent data reflecting distributions of both time specific residual errors and between-person intercept or slope variances. These data are then modeled using a set of linear equations (containing autoregressive, cross-lag, and slope loadings) in accordance with the respective population model to which they correspond to produce the observed variables to be used for analysis.

RI-CLPM: The error covariance matrix for the RI-CLPM is 14 X 14 having the form:

$$\Psi = \begin{bmatrix}
\sigma^2_{\alpha_{x_i}} & & & & & & & & & & & & & \\
\sigma_{\alpha_{x_i,y_i}} & \sigma^2_{\alpha_{y_i}} & & & & & & & & & & & & \\
0 & 0 & \sigma^2_{v_{x_1}} & & & & & & & & & & & \\
0 & 0 & 0 & \sigma^2_{v_{x2}} & & & & & & & & & & \\
0 & 0 & 0 & 0 & \sigma^2_{v_{x3}} & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{x4}} & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{x5}} & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{x6}} & & & & & & \\
0 & 0 & \sigma_{v_{x_1,y_1}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y1}} & & & & & \\
0 & 0 & 0 & \sigma_{v_{x_2,y_2}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y2}} & & & & \\
0 & 0 & 0 & 0 & \sigma_{v_{x_3,y_3}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y3}} & & & \\
0 & 0 & 0 & 0 & 0 & \sigma_{v_{x_4,y_4}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y4}} & & \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_{v_{x_5,y_5}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y5}} & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{v_{x_6,y_6}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{v_{y6}}
\end{bmatrix}$$

With mean structure:

$$\boldsymbol{\mu}^T = \begin{bmatrix} \mu_{\alpha_x} & \mu_{\alpha_y} & \mu_{x1} & \mu_{x2} & \mu_{x3} & \mu_{x4} & \mu_{x5} & \mu_{x6} & \mu_{y1} & \mu_{y2} & \mu_{y3} & \mu_{y4} & \mu_{y5} & \mu_{y6} \end{bmatrix}$$

The multivariate normal data generated from this error covariance matrix & mean structure contains the basic building blocks for building this model. The resulting variables consist of intercept terms, α_{Xi} & α_{Yi}, and time specific terms, ε_{Xit} & ε_{Yit}, which are each comprised of a mean component and variance component determined by the specifications given in PROC IML SAS 9.4. From here we generate sample means (μ_{Xt} & μ_{Yt}) and error variances (v_{Xit} & v_{Yit}) for each of the 1 through 6 time points to comprise the within-person component. With the between-person component being comprised of intercept means (μ_X & μ_Y) and variances (v_{Xi} & v_{Yi}), i.e.,

$$\varepsilon_{Xit} = \mu_{Xt} + v_{Xit} \qquad (9a)$$

$$\varepsilon_{Yit} = \mu_{Yt} + v_{Yit} \qquad (9b)$$

$$\alpha_{Xi} = \mu_X + v_{Xi} \qquad (9c)$$

$$\alpha_{Yi} = \mu_Y + v_{Yi} \qquad (9d)$$

In the RI-CLPM specifications we will be using for this study we'll have intercept means set to 0, with non-zero time varying means to capture change across time. With the sample data generated from the above procedure we model the within-person component by structuring residuals:

$$sr_{Xi1} = \varepsilon_{Xi1} \qquad (10a)$$

$$sr_{Yi1} = \varepsilon_{Yi1} \qquad (10b)$$

$$sr_{Xit} = \rho_{xx} sr_{Xi,t-1} + \rho_{xy} sr_{Yi,t-1} + \varepsilon_{Xit} \qquad (10c)$$

$$sr_{Yit} = \rho_{yy} sr_{Yi,t-1} + \rho_{yx} sr_{Xi,t-1} + \varepsilon_{Yit} \qquad (10d)$$

The final step in generating the observed variables to be used in analysis simply involves integrating the between- and within-person components:

$$x_{it} = sr_{Xit} + \alpha_{Xi} \qquad (11a)$$

$$y_{it} = sr_{Yit} + \alpha_{Yi} \qquad (11b)$$

These resulting variables are then exported as text files and compiled in a folder for analysis by MPlus 7.4.

LGCM-SR: The LGCM-SR data is generated in essentially the same manner as shown above for the RI-CLPM the only major difference is that we now include components associated with the random slope. Namely, we now have latent variables for the random slopes, each with their own means and variances, path coefficients to represent the loading of the observed variables on the slope, and the associated covariances amongst the slopes and intercepts. Under the LGCM-SR we now begin with a 16X16 error covariance matrix:

$$\Psi = \begin{bmatrix}
\sigma^2_{\alpha_{x_i}} \\
\sigma_{\alpha_{x_i},\alpha_{y_i}} & \sigma^2_{\alpha_{y_i}} \\
\sigma_{\alpha_{x_i},\beta_{x_i}} & \sigma_{\alpha_{y_i},\beta_{x_i}} & \sigma^2_{\beta_{x_i}} \\
\sigma_{\alpha_{x_i},\beta_{y_i}} & \sigma_{\alpha_{y_i},\beta_{y_i}} & \sigma_{\beta_{y_i},\beta_{x_i}} & \sigma^2_{\beta_{y_i}} \\
0 & 0 & 0 & 0 & \sigma^2_{x_1} \\
0 & 0 & 0 & 0 & 0 & \sigma^2_{x_2} \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{x_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{x_4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{x_5} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{x_6} \\
0 & 0 & 0 & 0 & \sigma_{x_1,y_1} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_1} \\
0 & 0 & 0 & 0 & 0 & \sigma_{x_2,y_2} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_2} \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_3,y_3} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_4,y_4} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_5,y_5} & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_5} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_6,y_6} & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{y_6}
\end{bmatrix}$$

With the mean structure

$$\mu^T = \begin{bmatrix} \mu_{\alpha_x} & \mu_{\alpha_y} & \mu_{\beta_x} & \mu_{\beta_y} & \mu_{x1} & \mu_{x2} & \mu_{x3} & \mu_{x4} & \mu_{x5} & \mu_{x6} & \mu_{y1} & \mu_{y2} & \mu_{y3} & \mu_{y4} & \mu_{y5} & \mu_{y6} \end{bmatrix}$$

As was done before we generate sample data from a multivariate normal distribution based on this error covariance matrix & mean structure. The resulting variables consist of intercept terms, α_{Xi} & α_{Yi}, slope terms, β_{Xi}& β_{Yi}, and time specific terms, ε_{Xit} & ε_{Yit} . All else is given as before,

with means and variances, in this case we add to the between-person component slope means ($\mu_{\beta x}$ & $\mu_{\beta y}$) and variances ($v_{\beta_{x_i}}$ & $v_{\beta_{x_i}}$), thus we now define our variables as,

$$\varepsilon_{Xit} = \mu_{Xt} + v_{Xit} \qquad (12a)$$

$$\varepsilon_{Yit} = \mu_{Yt} + v_{Yit} \qquad (12b)$$

$$\alpha_{Xi} = \mu_{\alpha_x} + v_{\alpha_{Xi}} \qquad (12c)$$

$$\alpha_{Yi} = \mu_{\alpha_y} + v_{\alpha_{Yi}} \qquad (12d)$$

$$\beta_{Xi} = \mu_{\beta_x} + v_{\beta_{Xi}} \qquad (12e)$$

$$\beta_{Xi} = \mu_{\beta_x} + v_{\beta_{Xi}} \qquad (12f)$$

In the case of the LGCM-SR we zero out the time specific means and allow the non-zero slope mean and variances to capture change over time. The intercept terms now have a non-zero mean to capture the mean at time point 1. As before we take the sample data generated from the error covariance and mean structure to model the within-person components by structuring residuals:

$$sr_{Xi1} = \varepsilon_{Xi1} \qquad (13a)$$

$$sr_{Yi1} = \varepsilon_{Yi1} \qquad (13b)$$

$$sr_{Xit} = \rho_{xx} sr_{Xi,t-1} + \rho_{xy} sr_{Yi,t-1} + \varepsilon_{Xit} \qquad (13c)$$

$$sr_{Yit} = \rho_{yy} sr_{Yi,t-1} + \rho_{yx} sr_{Xi,t-1} + \varepsilon_{Yit} \qquad (13d)$$

Like the procedure used with the RI-CLPM we integrate the between- and within-person components to produce our observed variables, except in this case we have the addition of the slope terms and their respective loadings:

$$x_{it} = sr_{Xit} + \alpha_{Xi} + \lambda_{Xt}\beta_{Xi} \qquad (14a)$$

$$y_{it} = sr_{Yit} + \alpha_{Yi} + \lambda_{Yt}\beta_{Yi} \qquad (14b)$$

The distinguishing feature between whether we produce the “unspecified” or linear trajectory is simply a matter of the lambda values we choose in equations 14a & 14b. As before the resulting dataset is exported as text files for analysis by MPlus 7.4.

3.3.1 Data Validation

The validation of the data generating process was conducted by fitting models to given generated sample data. Then the adequacy was evaluated by examining model fit to see that it indicates perfect to near perfect fit, and bias in parameter estimates was evaluated as well. Our primary index for bias of interest will be relative bias, $Relative\ Bias = \left(\frac{\hat{\theta}-\theta}{\theta}\right)$, and for model fit we will be primarily examining RMSEA, CFI, and χ^2 test for model fit.

RI-CLPM example: Utilizing the data generation procedure mentioned above and incorporating specific parameter values, we generated a specific data set as such:

Step 1: Set correlation values in PROC IML to reflect modest relationships between time specific errors and intercept variances, e.g., $Corr(\alpha_{Xi}, \alpha_{Yi}) = 0.30$ and $Corr(v_{Xit}, v_{Yit}) = 0.30$.

Step 2: Create a population covariance matrix utilizing variances from ECLS-K. Our vector of variances is given as:

$$\boldsymbol{V}^T = \begin{bmatrix} v_{\alpha_x} & v_{\alpha_y} & v_{x1} & v_{x2} & v_{x3} & v_{x4} & v_{x5} & v_{x6} & v_{y1} & v_{y2} & v_{y3} & v_{y4} & v_{y5} & v_{y6} \end{bmatrix}$$

With input values

$$\boldsymbol{V}^T = \begin{bmatrix} 141.5 & 106.9 & 103.94 & 213.5 & 213. & 213.5 & 213.5 & 213.5 & 81.99 & 134.03 & 134.03 & 134.03 & 134.03 & 134.03 \end{bmatrix}$$

We compute a covariance matrix using $\boldsymbol{D} = diag(\sqrt{\boldsymbol{V}})$ and the correlation matrix, $\boldsymbol{R}$, from step 1 as such: $\boldsymbol{S} = \boldsymbol{DRD}$. The resulting covariance matrix is given as:

$$\Psi = \begin{bmatrix} 141.5 \\ 36.9 & 106.9 \\ 0 & 0 & 103.94 \\ 0 & 0 & 0 & 213.5 \\ 0 & 0 & 0 & 0 & 213.5 \\ 0 & 0 & 0 & 0 & 0 & 213.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 213.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 213.5 \\ 0 & 0 & 27.69 & 0 & 0 & 0 & 0 & 0 & 81.99 \\ 0 & 0 & 0 & 50.75 & 0 & 0 & 0 & 0 & 0 & 134.03 \\ 0 & 0 & 0 & 0 & 50.75 & 0 & 0 & 0 & 0 & 0 & 134.03 \\ 0 & 0 & 0 & 0 & 0 & 50.75 & 0 & 0 & 0 & 0 & 0 & 134.03 \\ 0 & 0 & 0 & 0 & 0 & 0 & 50.75 & 0 & 0 & 0 & 0 & 0 & 134.03 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 50.75 & 0 & 0 & 0 & 0 & 0 & 134.03 \end{bmatrix}$$

Step 3: Utilizing the covariance matrix produced in step 2 and the mean vector derived from ECLS-K,

$$\mu^T = [\ 0 \quad 0 \quad 34.32 \quad 45.68 \quad 76.83 \quad 126.14 \quad 149.56 \quad 169.72 \quad 25.66 \quad 35.99 \quad 60.94 \quad 98.23 \quad 123.02 \quad 140.86\ ]$$

we go on to generate a sample data set of 1,000 observations by telling SAS to use a random multivariate normal distribution based on covariance matrix $\boldsymbol{S}$ and mean vector $\boldsymbol{\mu}$. The resulting dataset contains variables labeled Ix, Iy, Vx1-Vx6, and Vy1-Vy6 which respectively correspond to $\alpha_{Xi}, \alpha_{Yi}, \varepsilon_{Xi1} - \varepsilon_{Xi6}, \& \ \varepsilon_{Yi1} - \varepsilon_{Yi6}$ as given in the data generation description given above.

Step 4: Then setting the autoregressive parameters as $\rho_{xx} = \rho_{yy} = 1.20$ and the cross-lag parameters as $\rho_{yx} = \rho_{xy} = 0.80$ to reflect a situation where both constructs have the same level of stability with a reciprocal relation between the constructs. These values are applied to equation 10a-d then carried over into equations 11a-b to produce the observed variables, X1-

X6 and Y1-Y6. By using this procedure, it is necessary that we remove the mean values from the prior time points to prevent an exponential increase in the mean value of the scales across subsequent time points, this is what is represented in the subtractions below. The following set of equations represent our data generating code:

$$\varepsilon_{x_1} = v_{x_1}$$

$$\varepsilon_{y_1} = v_{y_1}$$

$$\varepsilon_{x_2} = 1.2(\varepsilon_{x_1} - 34.32) + 0.8\,(\varepsilon_{y_1} - 25.66) + v_{x_2}$$

$$\varepsilon_{y_2} = 1.2(\varepsilon_{y_1} - 25.66) + 0.8(\varepsilon_{x_1} - 34.32) + v_{y_2}$$

$$\varepsilon_{x_3} = 1.2(\varepsilon_{x_2} - 45.68) + 0.8\,(\varepsilon_{y_2} - 35.99) + v_{x_3}$$

$$\varepsilon_{y_3} = 1.2(\varepsilon_{y_2} - 35.99) + 0.8(\varepsilon_{x_2} - 45.68) + v_{y_3}$$

$$\varepsilon_{x_4} = 1.2(\varepsilon_{x_3} - 76.83) + 0.8\,(\varepsilon_{y_3} - 60.94) + v_{x_4}$$

$$\varepsilon_{y_4} = 1.2(\varepsilon_{y_3} - 60.94) + 0.8(\varepsilon_{x_3} - 76.83) + v_{y_4}$$

$$\varepsilon_{x_5} = 1.2(\varepsilon_{x_4} - 126.14) + 0.8\,(\varepsilon_{y_4} - 98.23) + v_{x_5}$$

$$\varepsilon_{y_5} = 1.2(\varepsilon_{y_4} - 98.23) + 0.8(\varepsilon_{x_4} - 126.14) + v_{y_5}$$

$$\varepsilon_{x_6} = 1.2(\varepsilon_{x_5} - 149.56) + 0.8\,(\varepsilon_{y_5} - 123.02) + v_{x_6}$$

$$\varepsilon_{y_6} = 1.2(\varepsilon_{y5} - 123.02) + 0.8(\varepsilon_{x_5} - 149.56) + v_{y_6}$$

Then using the residual terms (within person component) we create the observed variables by incorporating the between-person component:

$$x_1 = \varepsilon_{x_1} + \alpha_{x_i}$$

$$y_1 = \varepsilon_{y_1} + \alpha_{x_i}$$

$$x_2 = \varepsilon_{x_2} + \alpha_{x_i}$$

$$y_2 = \varepsilon_{y_2} + \alpha_{x_i}$$

$$x_3 = \varepsilon_{x_3} + \alpha_{x_i}$$

$$y_3 = \varepsilon_{y_3} + \alpha_{x_i}$$

$$x_4 = \varepsilon_{x_4} + \alpha_{x_i}$$

$$y_4 = \varepsilon_{y_4} + \alpha_{x_i}$$

$$x_5 = \varepsilon_{x_5} + \alpha_{x_i}$$

$$y_5 = \varepsilon_{y_5} + \alpha_{x_i}$$

$$x_6 = \varepsilon_{x_6} + \alpha_{x_i}$$

$$y_6 = \varepsilon_{y_6} + \alpha_{x_i}$$

Step 5: The data set produced via steps 1 through 4 is then exported as a text (.dat) file into a folder for subsequent analysis by MPlus 7.4. Within MPlus we specify a model that corresponds to the data generating model. In this case the between person components involve setting our random intercepts, zeroing out their means and freely estimating variances and covariances. We tell MPlus to estimate the time specific means in the observed variables. We set the observed variable variances to zero and fit time specific latent variables to capture the residual errors. Within each process, we constrain these residual variances to be equal across time but not across construct. The covariance between the x and y residuals at time 1 are freely estimated, but constrained as equal across subsequent time points. Autoregressive parameters within each construct are constrained to be equal across time lags. The cross-lag parameters from x to y are constrained equal across lags, and the cross-lag parameters from y to x are also constrained equal across lags, while, the x to y and y to x cross lag parameters are not constrained as equal.

Step 6: The resulting estimates from MPlus show, for the most part, decent recovery of the parameters from the population generating model. However, there are some cases where parameter recovery appears to be performing not so well. The relative bias within the intercept covariance and the residuals at time point one are high, about 17% overestimation in the intercept covariance and about 14% underestimation in the residual covariance at time point 1. To check the extent to which this poor recovery was driven by sample size, the generating sample was increased from 5,000 to 20,000, this resulted in the bias being reduced to 6%

underestimation of the intercept covariance and 3% overestimation for the time 1 residual covariance. Since our interest is not centrally focused on recovering these parameters, in the subsequent data generation, we will not use a generating sample of 20,000 to conserve computational requirements. The table below (table 4) shows model fit and relative bias in the respective parameter estimates:

Table 1. Results for Validating Data Generation from the RI-CLPM

PARAMS	TRUE	ESTIMATE	BIAS	Relative Bias
$Cov(\alpha_{Xi}, \alpha_{Yi})$	36.897	43.21	6.313	0.171098
$Var(\alpha_{Xi})$	141.5	148.863	7.363	0.052035
$Var(\alpha_{Yi})$	106.9	107.137	0.237	0.002217
$Var(\varepsilon_{Xi1})$	103.94	100.026	-3.914	0.037656
$Var(\varepsilon_{Xit}), t = 2, \ldots, 6$	213.5	204.189	-9.311	0.043611
$Var(\varepsilon_{Yi1})$	81.99	76.028	-5.962	0.072716
$Var(\varepsilon_{Yit}), t = 2, \ldots, 6$	134.03	130.547	-3.483	0.025987
$Cov\ (\varepsilon_{Xi1}, \varepsilon_{Yi1})$	27.694	23.76	-3.934	0.142052
$Cov\ (\varepsilon_{Xit}, \varepsilon_{Yit})$	50.748	48.202	-2.546	0.050169
ρ_{xx}	1.2	1.194	-0.006	0.005
ρ_{yy}	1.2	1.196	-0.004	0.003333
ρ_{xy}	0.8	0.81	0.01	0.0125
ρ_{yx}	0.8	0.805	0.005	0.00625
μ_{α_x}	0	0	0	0
μ_{α_y}	0	0	0	0
μ_{x_1}	34.32	33.955	-0.365	0.010635
μ_{x_2}	45.68	46.092	0.412	0.009019
μ_{x_3}	76.83	77.941	1.111	0.01446
μ_{x_4}	126.14	128.28	2.14	0.016965
μ_{x_5}	149.56	153.039	3.479	0.023262
μ_{x_6}	169.72	176.734	7.014	0.041327
μ_{y_1}	25.66	25.623	-0.037	0.001442
μ_{y_2}	35.99	35.383	-0.607	0.016866
μ_{y_3}	60.94	61.158	0.218	0.003577
μ_{y_4}	98.23	99.416	1.186	0.012074
μ_{y_5}	123.02	126.461	3.441	0.027971
μ_{y_6}	140.86	147.322	6.462	0.045875

RMSEA	5%	95%	CFI	SRMR
0.007	0	0.021	1.00	0.032
$\chi^2(65) = 68.48$, p=0.36				

The model fit, as expected, is good by all criteria. However, the RMSEA should be at zero and it is not.

LGCM-SR example: We proceed as before while incorporating the appropriate information for the LGCM-SR where we have now added means, variances and covariances associated with the slope terms. In the case of the linear trajectory, the slope loadings are not estimated as they will be with the unspecified growth trajectory. In contrast to the RI-CLPM, the time specific means are no longer estimated. Instead the intercept means are estimated in place of time 1 and the slope means are left to capture the change in the means over time.

Step 1: Set correlation values in PROC IML to reflect modest relationships between time specific errors, intercept variances, slope variances, and intercept-slope covariance within construct e.g., $Corr(v_{Xit}, v_{Yit}) = 0.30$, $Corr(\alpha_{Xi}, \alpha_{Yi}) = 0.30$, $Corr(\beta_{Xi}, \beta_{Yi}) = 0.30$, and, $Corr(\alpha_{Xi}, \beta_{Xi}) = 0.30$, while the cross construct intercept-slope covariance are set as being small in magnitude, $Corr(\beta_{Xi}, \beta_{Yi}) = 0.10$.

Step 2: Create a population covariance matrix utilizing variances from ECLS-K. Our vector of variances is given as:

$$\boldsymbol{V}^T = [\ v_{\alpha_x} \quad v_{\beta_x} \quad v_{\alpha_y} \quad v_{\beta_y} \quad v_{x1} \quad v_{x2} \quad v_{x3} \quad v_{x4} \quad v_{x5} \quad v_{x6} \quad v_{y1} \quad v_{y2} \quad v_{y3} \quad v_{y4} \quad v_{y5} \quad v_{y6}\ ]$$

With input values

$$\boldsymbol{V}^T = [\ 65 \quad 8 \quad 51 \quad 10 \quad 127 \quad 127 \quad 127 \quad 127 \quad 127 \quad 127 \quad 204 \quad 204 \quad 204 \quad 204 \quad 204 \quad 204\ ]$$

As before, we compute a covariance matrix using $\boldsymbol{D} = diag(\sqrt{V}\,)$ and the correlation matrix, $\boldsymbol{R}$, from step 1 as such: $\boldsymbol{S} = \boldsymbol{DRD}$. The resulting covariance matrix is given as:

$$\Psi = \begin{bmatrix}
65 &&&&&&&&&&&&&&& \\
6.84 & 8 &&&&&&&&&&&&&& \\
17.27 & 2.02 & 51 &&&&&&&&&&&&& \\
2.55 & 2.68 & 6.77 & 10 &&&&&&&&&&&& \\
0 & 0 & 0 & 0 & 127 &&&&&&&&&&& \\
0 & 0 & 0 & 0 & 0 & 127 &&&&&&&&&& \\
0 & 0 & 0 & 0 & 0 & 0 & 127 &&&&&&&&& \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 127 &&&&&&&& \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 127 &&&&&&& \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 127 &&&&&& \\
0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204 &&&&& \\
0 & 0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204 &&&& \\
0 & 0 & 0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204 &&& \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204 && \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 48.29 & 0 & 0 & 0 & 0 & 0 & 204
\end{bmatrix}$$

Step 3: Utilizing the covariance matrix produced in step 2 and a mean vector based on values derived from fitting an LGCM-SR to the ECLS-K data,

$$\mu^T = \begin{bmatrix} 37 & 22 & 48 & 26 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

we generate a sample data set of 5,000 observations by telling SAS to use a random multivariate normal distribution based on covariance matrix $\boldsymbol{S}$ and mean vector $\boldsymbol{\mu}$. The resulting dataset contains variables labeled Ix, Sx, Iy, Sy, Vx1-Vx6, and Vy1-Vy6 which respectively correspond to $\alpha_{Xi}\, \beta_{Xi}, \alpha_{Yi},\ \beta_{Xi}\,, \varepsilon_{Xi1} - \varepsilon_{Xi6}, \&\ \varepsilon_{Yi1} - \varepsilon_{Yi6}$ as given in the data generation description given above.

Step 4: Exactly as was done with the RI-CLPM we set the autoregressive parameters as $\rho_{xx} = \rho_{yy} = 1.20$ and the cross-lag parameters as $\rho_{yx} = \rho_{xy} = 0.80$. These values are applied to equations 13a-d then carried over into equations 14a-b to produce the observed variables, X1-X6 and Y1-Y6. The distinguishing feature between whether we are fitting the linear versus unspecified trajectory will be expressed in the slope loadings. This is represented by the following set of equations:

The Linear Trajectory

$$\varepsilon_{x_1} = v_{x_1}$$

$$\varepsilon_{y_1} = v_{y_1}$$

$$\varepsilon_{x_2} = 1.2(\varepsilon_{x_1}) + 0.8\,(\varepsilon_{y_1}) + v_{x_2}$$

$$\varepsilon_{y_2} = 1.2(\varepsilon_{y_1}) + 0.8(\varepsilon_{x_1}) + v_{y_2}$$

$$\varepsilon_{x_3} = 1.2(\varepsilon_{x_2}) + 0.8\,(\varepsilon_{y_2}) + v_{x_3}$$

$$\varepsilon_{y_3} = 1.2(\varepsilon_{y_2}) + 0.8(\varepsilon_{x_2}) + v_{y_3}$$

$$\varepsilon_{x_4} = 1.2(\varepsilon_{x_3}) + 0.8\,(\varepsilon_{y_3}) + v_{x_4}$$

$$\varepsilon_{y_4} = 1.2(\varepsilon_{y_3}) + 0.8(\varepsilon_{x_3}) + v_{y_4}$$

$$\varepsilon_{x_5} = 1.2(\varepsilon_{x_4}) + 0.8\,(\varepsilon_{y_4}) + v_{x_5}$$

$$\varepsilon_{y_5} = 1.2(\varepsilon_{y_4}) + 0.8(\varepsilon_{x_4}) + v_{y_5}$$

$$\varepsilon_{x_6} = 1.2(\varepsilon_{x_5}) + 0.8\,(\varepsilon_{y_5}) + v_{x_6}$$

$$\varepsilon_{y_6} = 1.2(\varepsilon_5) + 0.8(\varepsilon_{x_5}) + v_{y_6}$$

Then using the residual terms (within person component) we create the observed variables by incorporating the between-person component:

$$x_1 = \varepsilon_{x_1} + \alpha_{x_i}$$

$$y_1 = \varepsilon_{y_1} + \alpha_{x_i}$$

$$x_2 = \varepsilon_{x_2} + \alpha_{x_i} + \beta_{x_i}$$

$$y_2 = \varepsilon_{y_2} + \alpha_{x_i} + \beta_{y_i}$$

$$x_3 = \varepsilon_{x_3} + \alpha_{x_i} + 2\beta_{x_i}$$

$$y_3 = \varepsilon_{y_3} + \alpha_{x_i} + 2\beta_{y_i}$$

$$x_4 = \varepsilon_{x_4} + \alpha_{x_i} + 3\beta_{x_i}$$

$$y_4 = \varepsilon_{y_4} + \alpha_{x_i} + 3\beta_{y_i}$$

$$x_5 = \varepsilon_{x_5} + \alpha_{x_i} + 4\beta_{x_i}$$

$$y_5 = \varepsilon_{y_5} + \alpha_{x_i} + 4\beta_{y_i}$$

$$x_6 = \varepsilon_{x_6} + \alpha_{x_i} + 5\beta_{x_i}$$

$$y_6 = \varepsilon_{y_6} + \alpha_{x_i} + 5\beta_{y_i}$$

Step 5: As was done before we export the data as a text (.dat) file to analyze it in MPlus 7.4 by fitting the Linear LGCM-SR model to it. The between person components involve setting random intercepts, and random slopes with loading 0,1,2,3,4, &5 to reflect a linear growth trajectory in both constructs and based on the fitted model in MPlus we use the loadings λ=0,1,1.175,2.35,4.7,5 for the X construct and λ= 0, 1, 1.09,2.1875,4.37, 5 for the Y construct to make the non-linear trajectory. As can be seen the intercept is set at time 1 for both the linear and unspecified growth models.

The means, variances, and covariances amongst the intercept and slope terms are freely estimated. The time specific means and variances are zeroed out in the observed variables. The time specific residual errors are carried by fitting latent variables to each observed variable. Within each process, we constrain these residual variances to be equal across time but not across construct. The covariance between the x and y residuals are constrained to be equal across subsequent time as well.

Autoregressive parameters within each construct are constrained to be equal across time lags. The cross-lag parameters from x to y are constrained equal across lags, and the cross-lag parameters from y to x are also constrained equal across lags, while, the x to y and y to x cross lag parameters are not constrained as equal.

Step 6: As observed in the RI-CLPM recovery of random effects is rather poor. Relative bias in the variances and covariances in the intercept and slope terms is high in many cases. For example, the intercept-slope covariance for construct X is overestimated by about 44%, the intercept covariance between X and Y is nearly 20% overestimated, with the X slope- Y intercept covariance being underestimated by about 37%, the slope covariance between X and Y is not too bad at about 5% overestimation, the X intercept to Y slope covariance is overestimated at about 43%, and the intercept-slope covariance within the Y construct is not bad being less than 1%. The variances are tending to show underestimation for the most part ranging from about 9% to 17% underestimation. The Y intercept is overestimated by about 12%. In the within person component we have some underestimation as well, but it is low being around 2% underestimation for both the residual variances and covariances. As was done before, the reduction in bias due to increasing sample size from 5,000 to 20,000 can be evaluated to evidence that the poor random effect recovery is driven by sample size and not model misspecification. With the larger sample size the covariance biases are reduced: The X intercept-X slope covariance bias is now at 3.3%, X intercept-Y intercept covariance bias is now at -9.6%, X slope-Y intercept covariance bias is now at -15.2%, X slope-Y slope covariance bias is now at -8.5%, X intercept-Y slope covariance bias is now at 2.99%, and the Y intercept-Y slope covariance bias is reduced to -.71%. The variance biases were also reduced, but follow the same general pattern, with the Y intercept relative bias being overestimated only by 3.82%, and the remaining variance being underestimated from .5-3.6%.

The autoregressive, cross-lag, and mean estimates do not appear to be too concerning in terms of bias. All the model fit criteria are as we would want. The table below (table 5) shows model fit and relative bias in the respective parameter estimates for the linear trajectory.

Table 2. Results for Validating Data Generation from the Linear LGCM-SR

PARAMS	TRUE	ESTIMATE	BIAS	Relative Bias
$Cov(\alpha_{Xi}, \beta_{Xi})$	6.84	9.851	3.011	0.440205
$Cov(\alpha_{Xi}, \alpha_{Yi})$	17.27	20.406	3.136	0.181587
$Cov(\beta_{Xi}, \alpha_{Yi})$	2.02	1.282	-0.738	0.36535
$Cov(\beta_{Xi}, \beta_{Yi})$	2.68	2.802	0.122	0.045522
$Cov(\alpha_{Xi}, \beta_{Yi})$	2.55	3.648	1.098	0.430588
$Cov(\beta_{Yi}, \alpha_{Yi})$	6.77	6.794	0.024	0.003545
$Var(\alpha_{Xi})$	65	54.186	-10.814	0.16637
$Var(\beta_{Xi})$	8	6.846	-1.154	0.14425
$Var(\alpha_{Yi})$	51	57.044	6.044	0.11851
$Var(\beta_{Yi})$	10	9.148	-0.852	0.0852
$Var(\varepsilon_{Xit})$	127	124.22	-2.78	0.02189
$Var(\varepsilon_{Yit})$	204	198.21	-5.79	0.02838
$Cov\ (\varepsilon_{Xi1}, \varepsilon_{Yi1})$	48.29	47.161	-1.129	0.02338
$Cov\ (\varepsilon_{Xit}, \varepsilon_{Yit})$	48.29	47.161	-1.129	0.02338
ρ_{xx}	1.2	1.205	0.005	0.004167
ρ_{yy}	1.2	1.224	0.024	0.02
ρ_{xy}	0.8	0.8	0	0
ρ_{yx}	0.8	0.781	-0.019	0.02375
μ_{α_x}	37	36.867	-0.133	0.00359
μ_{β_x}	22	22.163	0.163	0.007409
μ_{α_y}	48	47.832	-0.168	0.0035
μ_{β_y}	26	25.758	-0.242	0.00931

RMSEA	5%	95%	CFI	SRMR
0.00	0	0.006	1.00	0.026
$\chi^2(68) = 51.684$, p=0.929				

For the unspecified growth curve, we have similar concerns as we saw with the linear curve. In this case, it seems we have a little less of a problem of bias in estimation amongst the latent intercept and slope variances and covariance, yet it is still problematic. In distinction to the linear model we now found more underestimation. The X slope-Y intercept covariance is underestimated by about 46%, with the X intercept-Y slope covariance being overestimated by about 42%, the Y slope and intercept covariance is underestimated by about 15%, the X-Y intercept covariance is overestimated by nearly 7%, and there is only about 2% underestimation in the slope to slope covariance. The intercept variance seems to have the higher bias with X intercept being underestimated by about 11% and the Y intercept being overestimated by about 19%, the slope variances are a bit lower with the X slope variance being overestimated by around 5%, and the Y slope variance is overestimated by about 6%. The residual errors are being more modestly misestimated, with the Y residuals underestimated by 4% and the X residuals by around 5%, the time one error covariance is overestimated by about 5% and at around 2% at subsequent time points. When the sample size is increased from 5,000 to 20,000 the covariance biases are reduced: X intercept-X slope reduces to 2.8%, the intercept covariance bias goes to 8%, X slope-Y intercept covariance bias reduces to 11%, X slope-Y slope covariance bias reduces to 6.23%, X intercept-Y slope covariance bias reduces to 12.9%, Y intercept-Y slope covariance bias reduces to 2.08%, and the residual covariances across time are essentially zero. In terms of variance, no bias exceeds 3%.

For the unspecified growth curve, we now see a bit more misestimation in the autoregressive and cross-lag parameters. The autoregressive weight for X are still below 1 %, the other cross-lag and autoregressive parameters now show a relative bias at around 2%. Apart from the Y slope mean which is underestimated at about 1.4% the other mean parameters all have relative bias less than 1%. The loadings were recovered well with all misestimation being less than 1%. Figure 7 shows the trajectories estimated overlay nearly perfectly with the true value trajectories. For the model fit indices, everything was good. The data validation results for LGCM-SR are given in table 3.

Table 3. Results for Validating Data Generation from the Unspecified LGCM-SR

PARAMS	TRUE	ESTIMATE	BIAS	Relative Bias
$Cov(\alpha_{Xi}, \beta_{Xi})$	6.84	8.01	1.17	0.171053
$Cov(\alpha_{Xi}, \alpha_{Yi})$	17.27	18.461	1.191	0.068964
$Cov(\beta_{Xi}, \alpha_{Yi})$	2.02	1.083	-0.937	0.463861
$Cov(\beta_{Xi}, \beta_{Yi})$	2.68	2.637	-0.043	0.016045
$Cov(\alpha_{Xi}, \beta_{Yi})$	2.55	3.615	1.065	0.417647
$Cov(\beta_{Yi}, \alpha_{Yi})$	6.77	5.722	-1.048	0.154801
$Var(\alpha_{Xi})$	65	57.889	-7.111	0.1094
$Var(\beta_{Xi})$	8	8.402	0.402	0.05025
$Var(\alpha_{Yi})$	51	60.829	9.829	0.192725
$Var(\beta_{Yi})$	10	9.354	-0.646	0.0646
$Var(\varepsilon_{Xit})$	127	122.376	-4.624	0.036409
$Var(\varepsilon_{Yit})$	204	194.335	-9.665	0.047377
$Cov\ (\varepsilon_{Xi1}, \varepsilon_{Yi1})$	48.29	50.48	2.19	0.045351
$Cov\ (\varepsilon_{Xit}, \varepsilon_{Yit})$	48.29	47.221	-1.069	0.022137
ρ_{xx}	1.2	1.193	-0.007	0.005833
ρ_{yy}	1.2	1.226	0.026	0.021667
ρ_{xy}	0.8	0.816	0.016	0.02
ρ_{yx}	0.8	0.783	-0.017	0.02125
μ_{α_x}	37	36.846	-0.154	0.004162
μ_{β_x}	22	22.191	0.191	0.008682
μ_{α_y}	48	47.829	-0.171	0.003562
μ_{β_y}	26	25.636	-0.364	0.014
λ_{X3}	4	4.03	-0.03	0.0075
λ_{X4}	8	8.06	-0.06	0.0075
λ_{X5}	16	16.11	-0.11	0.006875
λ_{X6}	17	17.09	-0.09	0.005294
λ_{Y3}	3	2.99	0.01	0.003333
λ_{Y4}	7	6.98	0.02	0.002857
λ_{Y5}	14	13.96	0.04	0.002857
λ_{Y6}	16	15.93	0.07	0.004375

RMSEA	5%	95%	CFI	SRMR
0.00	0	0.005	1.00	0.033
$\chi^2(60) = 47.379$, p=0.935				

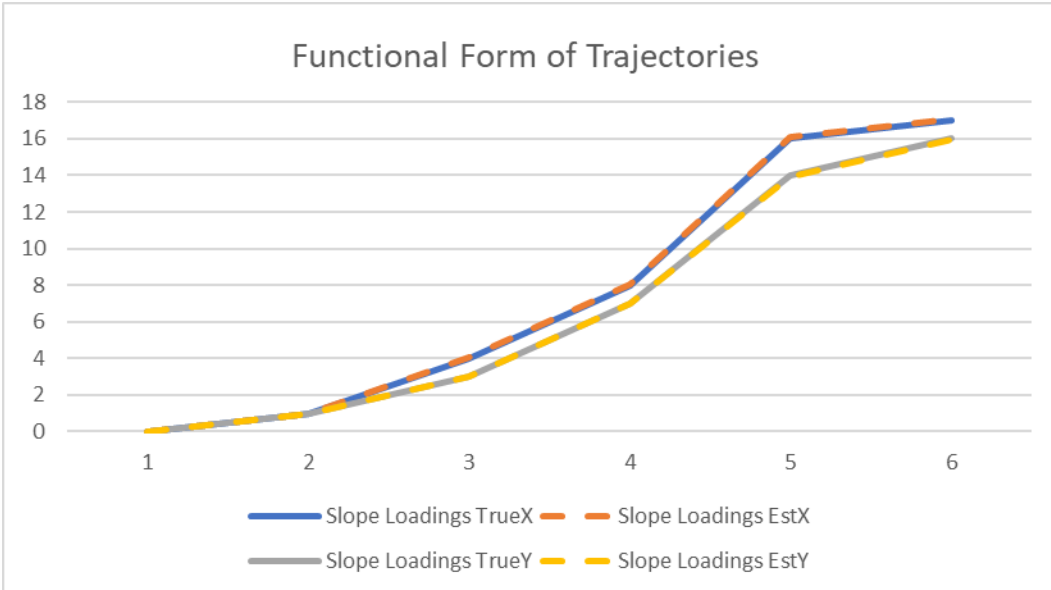


Figure 7. True and Estimated Trajectories from data validation

4.0 RESULTS

In the following we begin by presenting the results from the empirical study. Recall, in this study we were evaluating how different models could lead to different conclusions. The illustrative example used data from the Longitudinal Survey of American Youth for the 2007 to 2011 cohort (LSAY:2007-2011). The LSAY follows students from 7th to 12th grade to collect information about their development on various academic and non-academic characteristics. For our applied example, we focus on the relationship between Math Task Value (e.g., how valuable a student may think mathematical knowledge is) and Math Self-Concept (e.g., how capable with mathematics a student believes their self to be). Broadly speaking we are comparing amongst models that do not separate between and within-person effects (CLPM), models which account for stable between-person differences (RI-CLPM, FI-CLPM), and models which additionally account for inter-individual differences in change (LGCM-SR); the latter (LGCM-SR) models are further broken down into whether they allow for non-linear, unspecified growth (LGCM-SR-UGT) or only linear growth (LGCM-SR-LIN). The results of interest pertain to how well the different models fit and what conclusions we derive concerning whether Task Value dominates Self-Concept or vice versa depending on the model we fit.

Following the LSAY example, the results from the simulation study will be presented. As illustrated in detail within the methods section we have three models under consideration: the LGCM-SR-LIN, the LGCM-SR-UGT, and the RI-CLPM. The evaluation of these models involves generating data from a true model that is based on each of our three models, then subsequently fitting each of our three models to the data generated from the true models. This allows us to compare models relative to how well a model matching the true model performs. These results are organized by outcome, and within each outcome we consider what results under varying conditions of slope variance-covariance, slope loadings, and causal dominance conditions when fitting the different models to data which is either aligned with the true generating model or not.

The first outcome considered from the simulation study is the sheer number of times that admissible and convergent solutions were estimated. Consideration of when a model will produce valid estimates is a key consideration of the utility of a modeling approach. In this case, we consider the percentage of times that the solution produced by a model yield invalid results due to non-convergence, negative variance estimates, or correlations beyond an absolute value of one. Following the review of admissible and convergent solutions we consider model fit criteria. The central consideration regarding model fit criteria is the amount of time that an index correctly chooses the true model over the other models. Given that our conclusions about the relations amongst processes are based on the cross-lag parameters we consider these in three different ways. One, does a model capture the correct inference about the presence or absence and magnitude of a dominant process; two, to what extent does a model produce a biased estimate of a cross-lag path coefficient; and three, to what extent does a model produce bias in the standard errors for cross-lag path coefficients. When considering these components of our cross-lag paths, we will have a central interest in the relative performance of the different models in relation to the true model, as well as how this performance differs across the different slope variance-covariance conditions and whether a dominant process is present or not.

4.1 RESULTS FROM EMPIRICAL EXAMPLE WITH LSAY

Models were fit using robust maximum likelihood estimation methods. We assessed model fit by looking at information criteria (AIC, BIC), absolute fit indices (RMSEA, SRMR), and incremental/relative fit indices (CFI, TLI). Table 4 gives a complete listing of the model fit indices amongst the models. Because of scale differences we use standardized coefficient estimates to assess ARCL parameters. ARCL estimates are each averaged across times since these were fixed as constant across time. A complete listing of ARCL estimates are given in Table 5.

Table 4. Model Fit Amongst the Various Models

	AIC	BIC	RMSEA	SRMR	CFI	TLI
: CLPM	60043.43	60192.41	0.083	0.129	0.876	0.866
: CLPM(means)	60043.43	60192.41	0.083	0.129	0.883	0.871
: CLPM (Xit-Xi.)	1311469.00	131650.10	0.312	0.149	0.076	-0.002
: RI-: CLPM(means)	59613.62	59777.5	0.036	0.085	0.979	0.976
CM-SR						
: Linear	59558.98	59717.89	0.024	0.047	0.991	0.989
: UGT	59557.42	59756.06	0.023	0.035	0.992	0.990
: FI-CLPM						
/:Self-Concept)	126450.271	126728.220	0.088	0.063	0.809	0.753
/:Task Value)	127518.02	127795.97	0.087	0.057	0.801	0.743

Table 5. Autoregressive and Cross-Lag Estimates from the Various Models

	M1: CLPM	M2: CLPM (Xit-X.t)	M3:CLPM (Xit-Xi.)	M4: RI-CLPM	M5: LGCM- SR-Linear	M6: LGCM-SR-UGT	M7: FI-CLPM
TV8<-TV7	0.404	0.404	-0.150	0.247	0.167	0.158	0.167
TV9<-TV8	0.444	0.444	-0.150	0.266	0.184	0.171	0.168
TV10<-TV9	0.437	0.437	-0.150	0.252	0.175	0.163	0.162
TV11<-TV10	0.46	0.46	-0.150	0.277	0.197	0.185	0.163
TV12<-TV11	0.44	0.44	-0.150	0.248	0.174	0.158	0.165
SC8<-SC7	0.426	0.426	-0.010	0.138	0.087	0.076	0.211
SC9<-SC8	0.445	0.445	-0.010	0.14	0.088	0.077	0.223
SC10<-SC9	0.422	0.422	-0.010	0.126	0.08	0.07	0.225
SC11<-SC10	0.437	0.437	-0.010	0.131	0.084	0.073	0.224
SC12<-SC11	0.445	0.445	-0.010	0.137	0.089	0.076	0.218
TV8<-SC7	0.118*	0.118*	0.014	0.057	0.05	0.045	-0.056
SC8<TV7	0.103	0.103	-0.019*	0.09*	0.072*	0.067*	-0.058
TV9<SC8	0.124*	0.124*	0.014	0.057	0.05	0.045	-0.054
SC9<-TV8	0.113	0.113	-0.019*	0.097*	0.08*	0.072*	-0.060
TV10<-SC9	0.122*	0.122*	0.014	0.054	0.048	0.043	-0.054
SC10<-TV9	0.107	0.107	-0.019*	0.088*	0.072*	0.067*	-0.057
TV11<-SC10	0.133*	0.133*	0.014	0.062	0.056	0.05	-0.059
SC11<-TV10	0.107	0.107	-0.019*	0.088*	0.072*	0.066*	-0.057
TV12<-SC11	0.134*	0.134*	0.014	0.062	0.056	0.049	-0.062
SC12<-TV11	0.103	0.103	-0.019*	0.083*	0.068*	0.061*	-0.058

*indicates causally dominant path TV=task value, SC=self-concept

4.1.1 Results for CLPM Without Disentangled Effects

The CLPM on the observed variables (Model 1) did not yield particularly good model fit with RMSEA=0.083 (90% CI=0.076, 0.090), SRMR=0.129, CFI=0.876, and TLI=0.866. The autoregressive parameters for each self-concept and task value are relatively similar in magnitude averaging at 0.437 for task value and 0.435 for self-concept. The cross-lag estimates selected self-concept as causally dominating task value, showing on average that self-concept dominates task value by a factor of about 1.185. This is to say that self-concept->task value paths are nearly 20% stronger than the task value->self-concept paths. A test of path equality indicates that these paths are significantly different, Wald χ^2 (1) =19.119, p<0.001. All of the cross-lag parameters are significantly positive in sign, suggesting that individual' beliefs about their ability with mathematics leads to greater appraisal of the value of mathematics for them to a greater extent than increases in task value predict increases in self-concept. Time specific correlations between task value and self-concept have a correlation of about 0.5 at time 1 and around 0.4 at subsequent time points. Taken with the autoregressive estimates this implies that relationships between constructs at given time points is nearly as large as within construct relationships between time points.

Unsurprisingly, the CLPM with means (Model 2) had nearly equivalent fit to the CLPM on the observed variables, though by incremental fit indices we find it is a little more favored, with RMSEA=0.083 (90% CI=0.076, 0.090), SRMR=0.129, CFI=0.883, TLI=0.871. Furthermore, the autoregressive and cross-lag parameters are the same as found in the CLPM on observed variables, thus leading us to the exact same conclusion.

The CLPM fit directly to the person-mean centered variables (Model 3) was evaluated in terms of unstandardized coefficients. Since we are using centered variables, standardized results may be confusing. The fit was extremely poor, suggesting that perhaps this is not an acceptable approach, RMSEA=0.312 (90% CI=0.308, 0.316), SRMR=0.149, CFI=0.076, TLI=-0.002. Since we have fixed autoregressive and cross-lag parameters the values do not vary over time. Both task value and self-concept had negative autoregressions at -0.150 and -0.01 respectively. With this model, we now have task value dominating self-concept by a factor of about -1.357. A test of path equality indicated the difference in cross-lag estimates to be significant, Wald χ^2(1) =5.126, p=0.0236. The cross-lag suggest that when individuals are above their average on task value it will be followed by them below their average on self-concept. Likely, this effect is due to regression to the mean. The time specific correlations between constructs were as follows 0.467 at grade 7, 0.378 at grade 8, 0.412 at grade 9, 0.392 at grade 10, 0.428 at grade 11, and 0.439 at grade 12. Yet, due to such poor fit, these results are taken with caution.

4.1.2 Results from the RI-CLPM

The RI-CLPM (Model 4) improves in model fit over the CLPM, RMSEA=0.036 (90% CI=0.03, 0.04), SRMR=0.085, CFI=0.979, TLI=0.976. As one would anticipate, given that the random intercept is now absorbing some of the stability in processes, the autoregressive parameters reduce to about an average of 0.258 for task value and 0.134 for self-concept. This change from the prior models implies that variation accounted for by between person trait stability is likely greater for self-concept than task value. This can be further evaluated by examining the random intercept variances. Both task value and self-concept show significant interindividual variance in trait like stability. For task value, we have σ^2= 3.827, SE=0.324 yielding a test statistic of 3.827/0.324=11.822 which is significant at $p<0.001$. For self-concept, we have $\sigma^2 = 2.414$, SE=0.154 yielding at test statistic of 15.694 which is also significant at $p<0.001$. In terms of the cross-lagged parameters a new story emerges once between and within person processes are disentangled. The RI-CLPM indicates that task value is causally dominating self-concept by a factor of about 1.527, implying that the prediction of intra-individual changes in self-concept from intra-individual changes in task value is around 53% stronger than the prediction of task value from self-concept. A test of path equality indicates this difference to be significant, Wald $\chi^2(1)$ =6.818, $p=0.009$. All of the cross-lag estimates are significant and positive hence implying that increases in task value within a person predict increases in self-concept to a greater extent than the reverse. Note that, despite the poor fit, the person-centered CLPM yielded the same conclusion about causal dominance, except it determined it in a negative direction. The correlation between time specific residuals of the two processes is consistent with findings in the prior models having time 1 correlate around .426 and subsequent time points ranging from .35 to .41. Hence after accounting for trait like stability between persons, the time specific correlations between processes are greater than the within person-within process relation between time points. The covariance between random intercepts is significantly positive as well (2.054; correlation of 0.676), inferring that individuals who are generally higher on task value are higher on self-concept as well. This would appear to be consistent with findings from the cross-lags which indicate that increases in one process lead to increases in the other.

4.1.3 Results from the LGCM-SR

Since we are now including trajectories we begin this evaluation differently. First, we want to establish the shape of the trajectories for each process separately and then relate them together into LGCM-SR model. Figure 8 shows the resulting trajectories for task value when fit as linear vs. unspecified. Model fit indices do not, in general, show a clear preference for the linear vs. unspecified trajectory for task value (see Table 6). However, by looking at the unspecified trajectory it does appear that task value takes on a curvilinear growth trajectory. Figure 9 shows the resulting trajectories for self-concept. Self-concept also appears to take on a curvilinear growth trajectory, model fit is a bit more consistent in choosing the unspecified curve, except that BIC prefers linear curve, while both RMSEA and TLI show no preference (see Table 6).

Table 6. Model fit indices for Growth Curve Specifications

Growth Trajectories	AIC	BIC	RMSEA	SRMR	CFI	TLI
Self-Concept						
Linear	28727.28	28782.13	0.051	0.064	0.969	0.971
Unspecified	28716.71	28791.5	0.051	0.048	0.977	0.971
Task Value						
Linear	33483.17	33538.02	0.066	0.077	0.938	0.942
Unspecified	33464.08	33538.87	0.069	0.076	0.949	0.937

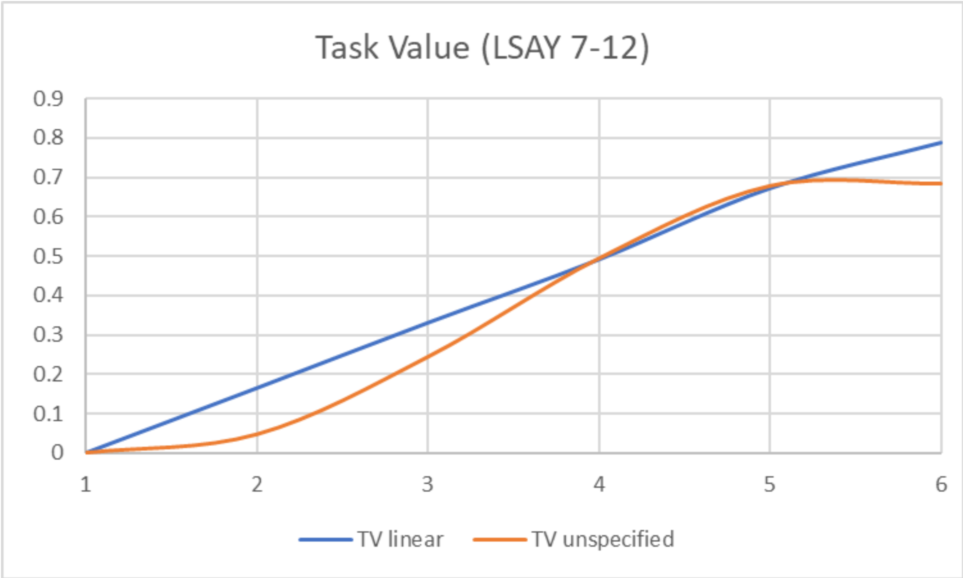


Figure 8. Growth Curves for Task Value

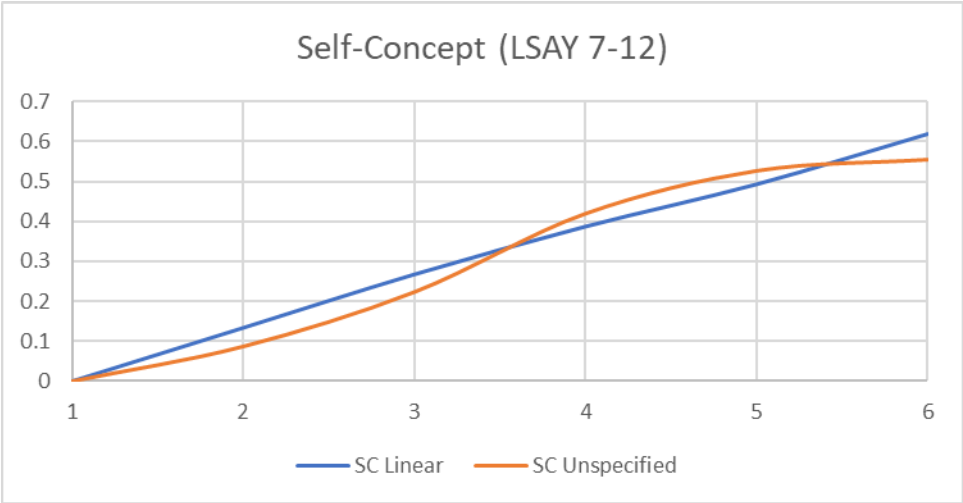


Figure 9. Growth Curves for Self-Concept

We first fit the bivariate model using the linear trajectory for both task value and self-concept (Model 5). The fit for this model was better than was observed with the RI-CLPM, RMSEA=0.024 (90% CI= 0.014, 0.032), SRMR=0.047, CFI=0.991, TLI= 0.989. The autoregressive parameters for task value makes a further drop to 0.179 and for self-concept to 0.086. As can be seen the added variance being accounted for in the between person's component further decreases the autoregressive parameters. For task value the intercept variance was 3.965, SE=0.709, yielding a test statistic of 5.593 which is significant at $p<0.001$ and the slope at 0.192, SE=0.054, yielding a test statistic of 3.52. Self-concept once again shows significant between person stability with 2.100, SE=0.272 (Wald's z=7.727, $p<0.001$), in terms of the slope we get the estimated variance 0.055, SE=0.021(Wald's z= 2.654, p=0.008). The cross-lag parameters indicate that task value is dominant. The cross-lag path coefficients show a bit of a decrease, with the factor by which task value dominates self-concept being reduced to 1.404. A test of path equality implies that there is no longer a significant difference between the cross-lag path coefficients, Wald $\chi^2(1)$ =0.687, p=0.4071. The covariance between task value and self-concept slopes is significant,0.484, SE=0.169 yielding test statistic=2.859 with p=0.004. This corresponds to a correlation value of 0.735, which indicates that growth in these two constructs is positively and closely related. The covariance between intercepts is also significant with (0.539, SE=0.072, test statistic=7.493, $p<0.001$). At time 1 the correlation in residual errors is 0.440, at later times they range from 0.34 to 0.43.

The resulting Latent Growth Curve Model with Unspecified Growth Trajectories (Model 6) showed near perfect fit with RMSEA=0.023 (90% CI=0.013, 0.033), SRMR=0.035, CFI=0.992, TLI=0.99, proving to be preferred over all the other models (see Table 4). The resulting trajectories are depicted in Figure 10, as one can see the trajectories take on a non-linear form. with a curvature that is not extreme, which gives some insight into why the unspecified trajectory presented only moderate improvement in the model fit over the linear trajectory. In terms of the autoregressive and cross-lag estimates we derive similar findings as with the Linear LGCM-SR. The coefficients undergo an even additional drop due to additional variance now being absorbed in the between person components by the unspecified growth curve, namely the inter-individual differences in change are being more deeply accounted for by allowing the curve to take on its own shape that best matches the data. The autoregressive estimates for task value averaged at 0.167 and those for self-concept at 0.074. Though all of the autoregressive estimates remain significant we do find an even greater drop in their value, and again self-concept is estimated even lower than task value autoregressive paths. For task value, intercept variance is estimated as $\sigma^2 = 3.987$, SE= 0.737 given a test statistic of 3.987/0.737= 5.410, $p<0.001$, and for self-concept $\sigma^2 = 2.117$, SE=0.264, 2.117/0.264= 8.019, $p<0.001$. For slope variances, we find that task value $\sigma^2= 4.233$, SE=1.488, giving a test statistic of 4.233/1.488=2.845, p=0.004, and self-concept has $\sigma^2 = 1.274$, SE=0.477, giving a test statistic of 1.274/0.477=2.671, p=0.008. Slope parameters are tricky to grasp within the unspecified framework, but what we can see is that there is comparably significant variation in interindividual differences in change for both task value and self-concept over time. This evidences that additional variance is being absorbed into the between person model from the within person model, which again is cohesive with our finding that there was reduction in autoregressive and cross-lag estimates.

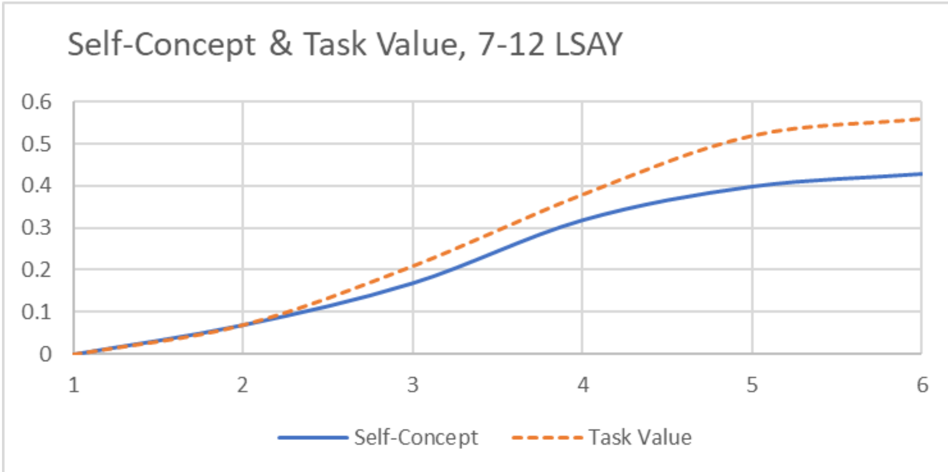


Figure 10. Unspecified Growth Curves for Self-Concept and Task Value

The LGCM-SR-UGT leads to the same conclusion about casual dominance as was given by the Linear LGCM & RI-CLPM, namely that task value dominates self-concept. Except in this case we now find that only the prediction of changes in self-concept from changes in task value is significantly positive, the self-concept->task value paths fall to non-significance. Task value is assessed as dominating self-concept by a factor of 1.442, approximately 44% stronger prediction of self-concept from task value than task value from self-concept. However, the test of path equality indicates that there is not a significant difference between the cross-lag path coefficients, Wald $\chi^2(1)$ =0.448, p=0.5032.

The decrease in the magnitude of cross-lag estimates between the RI-CLPM and LGCM-SR is notable. In prior models the cross-lag estimates were fairly consistent, whereas in the LGCM-SR estimates are reduced by about half (see Table 5) to see all of the cross-lag and autoregressive estimates from the various models. This sharp reduction may be related to the covariance of the task value and self-concept slopes pulling out some of the cross-lag effects over time. Examination of this covariance does show that there is a significant covariance between the slopes of the processes, $\sigma(_{TVslope,SCslope})$=1.292, SE=0.618, yielding test statistic 1.292/0.618= 2.091, p=0.037. Correlations between time specific residual are nearly the same as they were in RI-CLPM and Linear LGCM-SR, with some minimal reduction (7^{th} grade: 0.428, 8^{th} grade: 0.331, 9^{th} grade: 0.368, 10^{th} grade: 0.357, 11^{th} grade: 0.408, 12^{th} grade: 0.411). From these results, the addition of slope parameters has the greatest influence on the cross-lagged parameters as well creating some reduction in autoregressive and time specific residual correlations.

4.1.4 Results from the FI-CLPM

The FI-CLPM (Model 7) results show that when we treat self-concept as the lagged dv and task-value as the predicting variable our model fit is not very good, RMSEA=0.088 (90% CI=0.08, 0.094), SRMR=0.063, CFI=0.809, TLI=0.753. For self-concept, the FI-CLPM model presents a slightly higher standardized autoregressive parameter than the RI-CLPM, averaged as 0.165 across time. The intercept variance and standard error given by FI-CLPM for self-concept is a bit smaller than the estimate given by the RI-CLPM, $\sigma^2 = 2.037$, SE=0.125, yielding at test statistic of 16.355, $p<0.001$. All the correlations between the intercept and the time varying predictors are significant and increase across time, ranging from 0.333 at time 1 to 0.650 at time 6, the correlation between the lagged dv intercept and time 1 value is significant with a value of 0.485, and the time 1 correlation between task value and self-concept is significant with a value of 0.467.

When we treat task-value as the lagged dv predicted by self-concept, our model fit is similar to the converse model, RMSEA=0.087 (90% CI=0.081, 0.092), SRMR=0.057, CFI=0.801, TLI=0.743. For task-value the FI-CLPM model presents a slightly lower standardized autoregressive parameter than the RI-CLPM, averaged as 0.220 across time. The intercept variance and standard error given by FI-CLPM for task-value are again a bit smaller than the estimate given by the RI-CLPM, $\sigma^2 = 3.765$, SE=0.265, yielding at test statistic of 14.188, $p<0.001$. All of the correlations between the intercept and the time varying predictors are significant and increase across time, ranging from 0.388 at time 1 to 0.667 at time 6, the correlation between the lagged dv intercept and time 1 value is significant with a value of 0.399, and the time 1 correlation between task value and self-concept is significant with a value of 0.467, which is consistent with the model fit with self-concept as the lagged dv.

The cross-lag parameters derived from the FI-CLPM fit to each outcome present to be essentially equal, -0.058 when self-concept is the dependent variable and -0.057 when task-value is the dependent variable; the ratio is 1.02 indicating that task-value only dominates self-concept by about 2%. As can be seen here, the FI-CLPM, which is most comparable to the RI-CLPM amongst the models, shows some consistencies and some differences. Though estimates of the AR paths and variances are similar, we do find that variance estimates and their respective standard errors are smaller with the FI-CLPM. The change in AR parameters is not consistent, as in one case the FI-CLPM gives smaller path estimates and larger path estimates in the other case. The cross-lags reveal an entirely different story then all of the other models, implying that there is no clearly dominating process.

4.1.5 Summary of Results for LSAY Example

In general, what is found by model fit indices is that models which disentangle between and within person effects are preferred over ones that do not (see Table 4). The FI-CLPM is an exception to this. While the person mean centered CLPM is the worst overall, the FI-CLPM is second worst in terms of information criteria and relative fit indices. In terms of RMSEA the FI-CLPM is similar to the standard CLPM, though a bit poorer. However, SRMR indicates that FI-

CLPM is better than RI-CLPM but not as good as LGCM-SR. The LGCM-SR-UGT is considered the best fit of all the models, having what could be argued as excellent fit as compared to the other model. In Table 5 differences in the autoregressive and cross-lag estimates from each of the models are presented. Some of the key findings from this were that whether we disentangle between and within person effects or not could lead to different conclusions about which process is causally dominant. In this case, the conflated models lead to the conclusion that self-concept causally dominates task value whereas in the disentangled models we conclude that task value causally dominates self-concept, except in the case of the FI-CLPM where we conclude there is no real dominating process. It also seems that the addition of the random or fixed intercepts to account for trait like stability in processes amongst individuals is absorbing some of the autoregressive effects (aka stability parameters) from the within person models. Further, when slopes are fitted to the process cross-lag effects underwent a notable decrease in magnitude, implying that the slopes as well as their covariance were absorbing some of the within person cross-lagged effects. What is notable about the findings from the fixed-effect models is that they are not developed enough to give us good information about causal dominance since they do not test the reciprocal cross-lags at the same time.

4.2 CONVERGENCE RATES AND INADMISSIBLE SOLUTIONS

Issues in the correct fitting of a model came in three main forms: (1) negative variance estimates in the exogenous latent variables; (2) correlations amongst the exogenous latent variables that exceeded 1; and, (3) non-convergence. In the following the resulting datasets after removal of non-convergent and inadmissible solutions will be reported for each data set created from the

Monte Carlo simulations (see Table 7 for a summary of sample loss due to non-convergence and inadmissible solution).

When fitting the linear LGCM-SR model to the linear LGCM-SR data there was a total of 24 non-convergent solutions in the 54,000 replications. The more common issue was with estimating negative variances for the exogenous latent variables; there was a total of 4,458 inadmissible solutions due to this. In addition, there were a remaining 160 solutions that contained correlations amongst the exogenous latent variables that exceeded 1. The resulting sample, after removing these non-convergent and inadmissible solutions was N=49,358 (i.e., ~91.4% of the solutions were convergent and admissible, and thus were used in the analysis).

A likely reason underlying such non-convergence and these related inadmissible solutions is that the models being fit to the data are either overfitting the data (i.e., too much variance being explained due to fitting too many parameters) or down-fitting variance from the between person components to the within person components (i.e., misallocation of variance from one model component to another). Overfitting is suspected when fitting more complex models to data generated from simpler models: fitting LGCM-SR-UGT to RI-CLPM and LGCM-SR-Linear, and fitting LGCM-SR-Linear to RI-CLPM. While down-fitting is suspected when fitting simpler models to data generated from more complex models: fitting RI-CLPM to LGCM-SR-Linear and LGCM-SR-UGT, and fitting LGCM-SR-Linear to LGCM-SR-UGT.

As concerns the down-fitting piece, we see that when fitting simpler models to the more complex data between-person variance components are consistently underestimated while within-person residual errors are overestimated; further, this respective under- and over-estimation is highly correlated. To illustrate, when fitting the RI-CLPM to the Linear LGCM-SR, we have the X intercept underestimated by -63.68% and the Y intercept underestimated by -

152.60%, with the X residual (t=1) being overestimated by +37.96% and the Y residual (t=1) being overestimated by +43.78%. The correlations between the intercepts and residuals (t=1) all exceed r=-0.99, suggest an essentially perfect negative relationship wherein as overestimation of the residual variance estimates increase so too does the underestimation of the intercept variance estimates. Similarly, when fitting the RI-CLPM to the LGCM-SR-UGT we have residual errors being overestimated between +21% and +22% and intercept variance being underestimated between -39.71% for the X intercept variance and -75.84% for the Y intercept variance. The correlations amongst these misestimations are again high, with all the intercept variance biases and the residual variance (t=1) biases exceeding r=-0.97, suggesting again a near perfect correspondence wherein as the overestimation of the residual variance goes up so too does the underestimation in the intercept variance. When fitting the Linear to the Unspecified LGCM-SR the underestimation of the intercepts and the overestimation of the residual errors is severe (note: in this condition we had no admissible solutions). The X intercept variance was underestimated by -1572.45% and the Y intercept variance was underestimated by -980.53%, while the Y residuals were overestimated by +442.72% and the X residuals were overestimated by +566.13%. However, the correlations were less extreme, but still notable, ranging between r = -0.269 for X residual overestimation to Y intercept variance underestimation, and, r = -0.785 for Y residual overestimation to X intercept variance underestimation. Again indicating that as underestimation in the between-person variance components increases so too does the overestimation in the within-person variance components. Regarding overfitting more complex models to data generated from simpler models, we find that anytime we fit LGCM-SR Linear and UGT to the RI-CLPM slope variance estimates are also negative, clearly indicating an overfitting due to the fact that RI-CLPM contains no inter-individual variance in change over time and the LGCM-SR models are trying to estimate it anyhow. When fitting the LGCM-SR-UGT to the LGCM-SR-Linear, we see that in nearly half the cases (45.58%) our relative fit indices are exceeding 1 which signals overly perfect modeling of the Linear LGCM-SR data by the LGCM-SR-UGT. In sum, it is plausible that complex models are overfitting data from simpler models, and simpler models are misallocating error between the within- and between-person level (down-fitting), which may be underlying our admissible and convergent solution issues.

Considering non-convergent and inadmissible solutions when fitting the Unspecified LGCM-SR model to the Linear LGCM-SR data, one case was lost due to non-convergence, 3,079 were dropped due to negative variance estimates, and an additional 3,929 were dropped for having correlations amongst the exogenous latent variables greater than 1 or less than -1. This results in a set of 46,991 cases, for a convergence rate of 87%. Fitting the RI-CLPM to the Linear LGCM-SR data was considerably more problematic. With a total of 1,622 non-convergent solutions, a total of 11,822 solutions were inadmissible due to negative variance estimates, and a total of 5,138 solutions were inadmissible due to correlations amongst the latent variables being outside the 1 to -1 range. This means that only about 65% (N=35,418) of the results yielded from fitting the RI-CLPM to the Linear LGCM-SR were used in analysis. Again, the main issue was due to the non-dominant condition. However, in this case it does appear that increasing variance and covariance in the slopes was also negatively impacting the results in terms of convergent and admissible solutions.

When generating from the Unspecified LGCM-SR, fitting the generating model was less problematic in terms of non-convergent and inadmissible solutions. Only 1 replication didn't converge, there were no negative variance estimates, and only 2 correlations in the latent variables were out of range. Thus, the resulting sample was 53,997, which is negligible (convergence rate of 99.9%). When fitting the Linear LGCM-SR to the Unspecified LGCM-SR data convergence and inadmissible solutions were extremely problematic. There were 647 non-convergent solutions. The biggest issue was with inadmissible solutions; 53,224 replications produced negative variance estimates. This results in only 129 observations, however, amongst these remaining cases the correlation between the first-time point was less than negative 1, thus in no case did the Linear LGCM-SR produce an admissible or convergent solution. In this way, we can argue that in terms of our standards, the Linear LGCM-SR should not be used in the presence of a non-linear trajectory.

In terms of producing valid solutions, the RI-CLPM proves to be a better option than a linear model when fitting to a non-linear trajectory. The RI-CLPM only produced 9 non-convergent solutions, about 20% of the solutions (N=10,893) were inadmissible due to negative variance estimates, and around 9% were inadmissible due to correlations outside of the 1 to -1 range (N=4,778), for a remaining analytic sample of 38,320 (~71%).

For the RI-CLPM generating model we only have 2,000 cases and when we fit the RI-CLPM to this data all replicated results were convergent and admissible. When fitting the Linear LGCM-SR to the RI-CLPM there were no convergence issues, but 100% of the models produced negative variance estimates in the latent variables, namely the X slope is always estimated with a negative variance. This is likely because the analysis model contains more variance components than are produced by the generating model. A similar thing happens when fitting the Unspecified LGCM-SR, wherein 296 of the models do not converge, and in all cases, a negative variance for the latent variables was produced, namely the X slope always has a negative variance.

Additionally, unbounded correlations were present, along with other negative variance estimates, such that we never had an admissible solution.

Table 7. Summary of Convergent and Admissible Solutions

		Remaining Datasets after loss due to:			
Generating Model	Fitted Model	Non-Convergence	Negative Variance	Correlations > 1	Convergent and Admissible Solutions
LGCM-SR-LIN	LGCM-SR-LIN	53,976	49,518	49,358	91.4%
Total=54,000	LGCM-SR-UGT	53,999	50,920	46,991	87%
	RI-CLPM	52,378	40,556	35,418	65.6%
LGCM-SR-UGT	LGCM-SR-UGT	53,999	53,999	53,997	99.99%
Total=54,000	LGCM-SR-LIN	53,353	129	0	0%
	RI-CLPM	53,991	43,098	38,320	70.96%
RI-CLPM	RI-CLPM	2,000	2,000	2,000	100%
Total=2,000	LGCM-SR-LIN	2,000	0	0	0%
	LGCM-SR-UGT	2,000	1,704	0	0%

4.3 MODEL FIT

When comparing models in terms of fit we look at this in accordance with three classes of model fit indices: information criteria, relative fit, and absolute fit. Information criteria must be considered relative between models as the numbers themselves are not readily interpretable in terms of a null hypothesis or absolute goodness of fit. If all the models being compared fit poorly we will have no sense of this from the information criteria alone, so rather we use it to compare between models. The general rule of thumb is that a more minimal information criteria indicates fit preference for the model yielding the smaller value. Details on the information criteria being assessed in this paper can be found in the methods section on dependent variables used in the simulation study.

The relative fit indices between models are also considered comparatively between models, and are evaluated in relation to a null model. Unlike information criteria we do have criteria values to evaluate model fit, with a value of 1.00 indicating a perfectly fit model. The absolute model fit indices are also compared between models, with a value of 0.00 indicating a perfectly fit model, these indices are called absolute because they make a comparison directly to a saturated (i.e., fully explained) model.

When considering our fit indices, our main interest concerns how often the index favors the correct model as this aligns with the way in which model fit indices are used in practice. In the case that a model index doesn't indicate a preference, we opt for parsimony. For example, if the same index value were derived for the RI-CLPM and the Linear LGCM-SR, then we say that the index favors the RI-CLPM because it is a less complex model. Because of issues with non-convergence and inadmissible solutions barring us from comparing between models in given replications, we treat such cases as missing values. In the following we present the percentages for the number of replications in which the various indices select the correct model. Table 8 summarizes correct selection rates across the model fit indices.

4.3.1 Information Criteria Results

When fitting to the Linear LGCM-SR data, overall, the AIC did well at selecting the correct model over the Unspecified LGCM-SR, for a total of 96% correct selection, and slightly less well at selecting the correct model over the RI-CLPM, for a total of 91% correct selection. The slope variances were an important factor in making the correct selection over the RI-CLPM, namely as slope variances increased the AIC became more likely to correctly choose the Linear LGCM-SR over the RI-CLPM. This result would be expected when considering that the RI-CLPM does not account for slope variance as the LGCM-SR models do. Specifically, the percentage of time that AIC correctly chooses over the RI-CLPM goes from 84% to 100% when moving from an X slope variance of 4 to 16, and 88% to 95% from Y slope variance of 5 to 20. The slope covariance also appeared to play some role in selecting the correct model, moving from 94% correct selection over the RI-CLPM when the correlation was small to 90% when the correlation was large. However, the most important factor pertained to the dominance condition, with the AIC only selecting the correct model 71% of the time when dominance was present as opposed to 99% of the time when no dominance condition was present. When fitting RI-CLPM to the Unspecified LGCM-SR data, we find that AIC selects the correct model 100% of the time.

The BIC performed similarly; however, due to its more strenuous penalty for model complexity, we find that the BIC selects the correct Linear LGCM-SR over the Unspecified LGCM-SR 100% of the time. Overall, the BIC selects the correct, Linear LGCM-SR, model over the RI-CLPM 92% of the time. Again, slope variances impact the correct selection over the RI-CLPM, going from 84% correct selection at low X slope variance to 100% correct selection at the highest X slope variance and 89% correct selection at low Y slope variance to 95% correct selection at high Y slope variance. At the low slope covariance, the BIC correctly chooses over the RI-CLPM 94% of the time and 90% of the time when the slope covariance is high. The dominance condition, as with AIC, appears to be the most influential factor on correct selection over the RI-CLPM using BIC with 72% correct selection when a dominant process was present to 99% when there was no dominant process. When fitting to the Unspecified LGCM-SR data, BIC correctly selected over RI-CLPM 100% of the time.

4.3.2 Relative Fit Indices

The Comparative Fit Index (CFI) performed well at correctly choosing the Linear LGCM-SR over the RI-CLPM 100% of the time, but not as well at correctly choosing over the Unspecified

LGCM-SR (a rate of 74% correct selection). There was no clear factor driving the correct selection as the rate was consistent across conditions. In fact, what was most salient when considering correct selection rates between the Linear and Unspecified LGCM-SR was that the CFI value was always at or extremely near one implying perfect fit for both models. Thus, the primary reason that the CFI is selecting the correct, Linear LGCM-SR model over the Unspecified LGCM-SR was likely due to it being a less complex model. From this CFI is not particularly useful when considering potentially saturated models. When fitting the RI-CLPM to the Unspecified LGCM-SR data, CFI correctly selected the Unspecified LGCM-SR over RI-CLPM 100% of the time.

Similarly, Tucker-Lewis Index (TLI) correctly selected Linear LGCM-SR over RI-CLPM 100% of the time, but only correctly selected over the Unspecified LGCM-SR 55% of the time. As was the case with CFI no clear pattern of variation in correct selection rates across conditions was apparent with both Unspecified and Linear LGCM-SR producing TLI values at or near one and Linear being selected because it was less complex. The Unspecified LGCM-SR was correctly selected over the RI-CLPM 100% of the time by TLI. On average, the differences in relative fit index values between the Linear and Unspecified LGCM-SR when fitting to the Linear LGCM-SR data was less than one-ten-thousandth.

4.3.3 Absolute Fit Indices

When fitting to the Linear LGCM-SR data, we find that RMSEA chose the correct model over the RI-CLPM 100% of the time and over the Unspecified LGCM-SR 79% of the time. No apparent pattern of model selection differences across conditions emerged. Similar to the case with the relative fit indices we have index values indicating perfect fit for both LGCM-SR models, thus correct selection was likely made because of the Linear model being more parsimonious. As was the case with the relative fit indices, the average difference between RMSEA for Linear vs. Unspecified LGCM-SR when fitting to the Linear LGCM-SR was less than one-ten-thousandth of a point. The RMSEA correctly chose the Unspecified LGCM-SR over the RI-CLPM 100% of the time.

SRMR did well at correctly choosing the Linear LGCM-SR over the RI-CLPM 99% of the time but poorly at correctly choosing over the Unspecified LGCM-SR, with a correct selection rate of 33%. There appeared to be a minor effect of dominance condition on the correct selection of Linear LGCM-SR over the Unspecified LGCM-SR, with 37% correct selection when no dominant process was present and 31% in the presence of a dominant process. The likely reason behind this finding may be that the SRMR does not account for any penalty due to model complexity and is purely founded on which model explains the most variance. Because the unspecified model freely fits a trajectory, it can account for more sample variance than the linear model, thus consistently producing a lower SRMR value overall. The SRMR correctly selected the Unspecified LGCM-SR over the RI-CLPM 99% of the time. The dominance condition and X slope variance may have some minor influence on the correct selection of the Unspecified LGCM-SR over the RI-CLPM, with 98% correct selection when there was a dominant process and nearly 100% when there wasn't a dominant process, and we move from 98% to 100% going from low to high X slope variance.

Table 8. Correct Selection Rates amongst the model fit indices

<table>
<tr><td></td><td colspan="2">Fitting to Linear LGCM-SR</td><td colspan="2">Fitting to Unspecified LGCM-SR</td></tr>
<tr><td>Model Fit Index</td><td>LGCM-SR-UGT</td><td>RI-CLPM</td><td rowspan="7">Recall LGCM-SR-LIN never produces an admissible solution when fit to LGCM-SR-UGT</td><td>RI-CLPM</td></tr>
<tr><td>AIC</td><td>96%</td><td>91%</td><td>100%</td></tr>
<tr><td>BIC</td><td>100%</td><td>92%</td><td>100%</td></tr>
<tr><td>CFI</td><td>74%</td><td>100%</td><td>100%</td></tr>
<tr><td>TLI</td><td>55%</td><td>100%</td><td>100%</td></tr>
<tr><td>RMSEA</td><td>79%</td><td>100%</td><td>100%</td></tr>
<tr><td>SRMR</td><td>33%</td><td>99%</td><td>99%</td></tr>
</table>

4.4 CROSS-LAG PARAMETERS

In the following we consider the effect of the models and conditions on the cross-lag components. This is important as these are the components from which our inferences and conclusions about phenomena under study with such models will be based. We consider these in three different ways: (1) in terms of the dominance factor; (2) in terms of cross-lag parameter estimate bias; and, (3) bias in the standard errors for the cross-lag parameter estimates. The factor of dominance is of a more substantive interest as it reflects the conclusions we make about causally dominant processes. The methods section gives more detail on how this measure is formulated. The relative bias in the cross-lag parameter estimates is of interest because it gives us insight into the conditions under which parameter estimation is more or less trustworthy. The corresponding standard error relative bias is important as it gives insight into the confidence with which we make these estimates and the extent to which the statistical testing correctly captures an effect. When considering these findings we encounter higher order interactions, in the presence of such interaction lower order and main effects should be interpreted with caution. In the presentation of these findings with higher order interactions, lower order interactions and main effects are presented as accessories for understanding higher order interaction.

Because of the large sample being used in these analyses we will be less concerned with statistical significance as gauged by p-values which are sample size dependent, and will rather focus on effect sizes. Specifically, we will use partial eta-squared, $\eta_p^2 = \frac{SS_{effect}}{SS_{effect}+SS_{error}}$, which is a measure of the proportion of variance explained by the models and conditions under evaluation after controlling for one another. As a criteria value, we will use a partial η^2 value of 0.10 to signal what we can consider to be notable influences, as this would correspond to 10% of partial variance explained.

4.4.1 Factor of Dominance

The factor of dominance can be evaluated under the understanding that when Y is dominating X the factor should be 4, whereas when there is no dominant process the factor should equal one. When fitting to the Linear LGCM-SR data, naturally the most important factor for factor of dominance was the dominance condition itself, $\eta_p^2 = 0.9816$. There was also a significant interaction between the model and the X slope variance, $\eta_p^2 = 0.2008$, this effect is reflected in table 9.

Table 9. Factor of dominance when fitting the LGCM-SR data by dominance condition, X slope variance, and model

Dominance Condition	Model	X slope variance	N	Mean	SD
Y dominates X	Linear LGCM-SR	4	8938	4.003	0.106
		8	8936	4.002	0.106
		16	8921	4.001	0.109
	Unspecified LGCM-SR	4	8942	4.003	0.106
		8	8937	4.002	0.106
		16	8922	4.001	0.109
	RI-CLPM	4	6596	3.932	0.235
		8	1952	3.520	0.209
No Dominance	Linear LGCM-SR	4	7516	1.000	0.014
		8	7554	1.000	0.015
		16	7493	1.000	0.015
	Unspecified LGCM-SR	4	8997	1.000	0.014
		8	6846	0.999	0.015
		16	4347	1.000	0.015
	RI-CLPM	4	9000	1.134	0.145
		8	8997	1.038	0.132
		16	8873	0.875	0.113

Overall, there was high consistency in accurately capturing the proper factor of dominance when fitting the LGCM-SR models to the Linear LGCM-SR data. When fitting RI-CLPM to a Linear LGCM-SR with a dominant process, increases in the X slope variance led to greater underestimation of the dominance relation and when the X slope variance reached its highest level we failed to estimate the models at all. When no dominant process was present it was not entirely clear what the pattern was, when moving from an X slope variance of 4 to an X slope variance of 8, the capturing of this dominance factor appeared to improve from being overestimated to being more just estimated. However, when going from X slope variance of 8 to X slope variance of 16 we saw a worsening in the adequacy of capturing the dominance relation, such that underestimation of this factor occurred.

As expected the largest effect on the factor of dominance pertained to the dominance condition itself, $\eta_p^2 = 0.9512$, but this finding was trivial. Of more interest are the effects of the model being fit to the LGCM-SR-UGT data, $\eta_p^2 = 0.1932$, particularly as moderated by the dominance condition, $\eta_p^2 = 0.1902$. The RI-CLPM only slightly underestimated the factor of dominance, implying that it performs well when fit to the LGCM-SR-UGT data. Table 10 shows this effect, when fit to itself the factor of dominance is precisely captured, while there was some systematic underestimation on behalf of the RI-CLPM model being fit to the LGCM-SR-UGT data.

Table 10. Factor of Dominance between models across dominance conditions when fit to the LGCM-SR-UGT data

Model	Dominance Condition	N	M	SD
LGCM-SR-UGT	X Dominates Y	26998	4.000	0.090
	No Dominance	26999	1.000	0.014
RI-CLPM	X Dominates Y	11331	3.640	0.287
	No Dominance	26989	0.991	0.159

Further, we found that the X slope variance influenced how well the factor of dominance was captured by the RI-CLPM model being fit to the LGCM-SR-UGT data: $\eta_p^2 = 0.2171$ for the three-way interaction model by dominance condition by X slope variance, and $\eta_p^2 = 0.4110$ for the two-way model by X slope variance interaction. In Table 11 we can see that within the dominant process condition increasing the X slope variance led to greater underestimation of the dominance factor, and perhaps more notably within the highest variance condition we only have 9 valid observations, meaning that this increasing variance is leading to more inadmissible or non-convergent solutions. The pattern in the non-dominance condition was less clear, where it began by slightly overestimating and then within the highest X slope variance condition became notably underestimated.

Table 11. Factor of dominance across X slope variance and dominance condition when fitting RI-CLPM to the LGCM-SR-UGT

Model	Dominance Condition	X Slope Variance	N	Mean	SD
RI-CLPM	X Dominate Y	4	7395	3.797	0.195
		8	3927	3.347	0.176
		16	9	2.762	0.146
	No Dominance	4	9000	1.112	0.129
		8	9000	1.012	0.118
		16	8989	0.849	0.101

Additionally, we find an effect of the Y slope variance on capturing the factor of dominance, $\eta_p^2 = 0.3045$. In table 12, though we do not have a significant three-way interaction, we depict the dominance conditions and Y slope variance levels for interpretability. It appears that when there was a dominant process there was an improvement in estimation as the Y slope variance increased, though admissible or convergent solutions appeared to degrade. The non-dominant condition indicated that underestimation became less severe moving from low to medium Y slope variance, but, became overestimated when going from medium to high variance. Figure 11 illustrates the influence of X and Y slope variance across dominance conditions when fitting RI-CLPM to LGCM-SR-UGT.

Table 12. Factor of dominance across Y slope variance and dominance condition when fitting RI-CLPM to the LGCM-SR-UGT

Model	Dominance Condition	Y Slope Variance	N	Mean	SD
RI-CLPM	X Dominates Y	5	4488	3.485	0.222
		10	4104	3.626	0.231
		20	2739	3.916	0.251
	No Dominance	5	9000	0.872	0.098
		10	9000	0.955	0.105
		20	8989	1.146	0.127

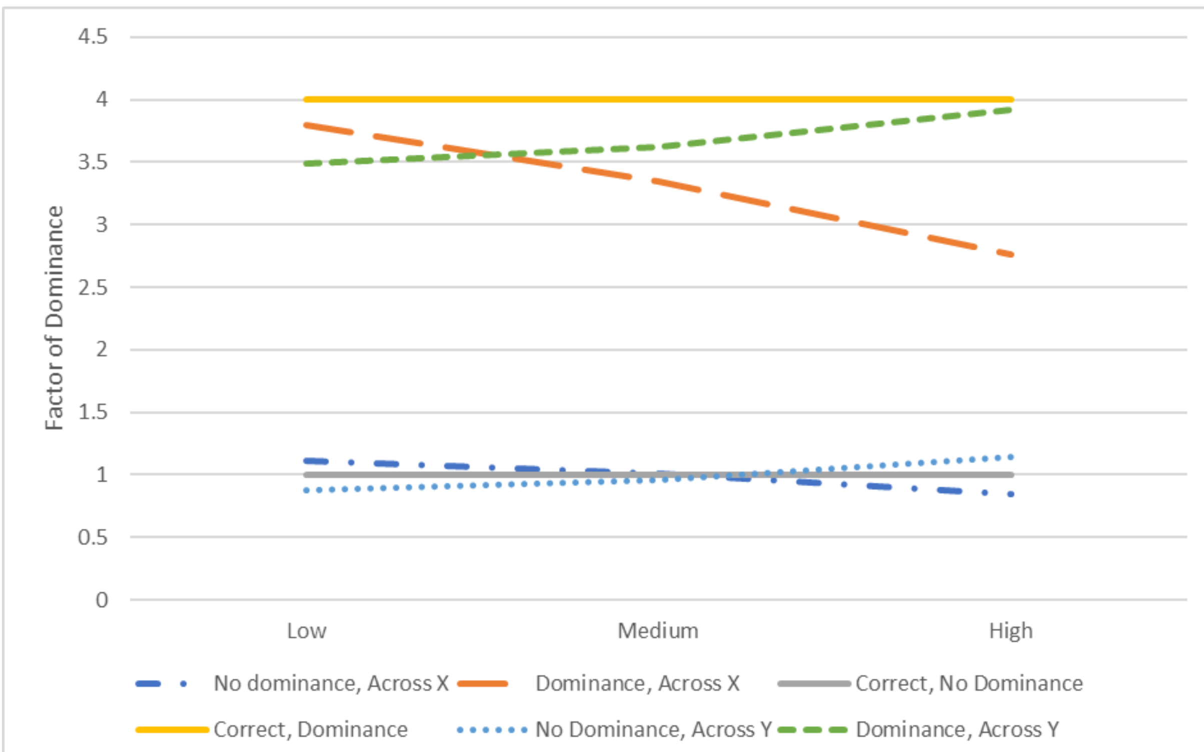


Figure 11. Factor of Dominance across dominance condition and slope variance

In sum, the major finding with the factor of dominance is that when fitting the RI-CLPM to the Unspecified LGCM-SR data we have notable underestimation of the factor of dominance especially in the presence of a dominant process. Interestingly, the underestimation becomes more severe as the X slope variance increases (i.e., dominant process variability), but becomes less severe as the Y slope variance increases (i.e., dominated process variability).

4.4.2 Relative Bias in the Cross-Lag Parameter Estimates

When fitting to the Linear LGCM-SR data we found an interaction effect on relative bias in the cross-lag parameter estimates regarding a predictor's slope variance and the model being fit to the data. Table 13 gives some description of this effect. For the relative bias in the Y to X pathways we found a main effect for Y slope variance, $\eta_p^2 = 0.1839$, model, $\eta_p^2 = 0.5170$, and an interaction between the model and Y variance, $\eta_p^2 = 0.1956$. A closer examination of this effect reveals that relative bias in the cross-lag parameter estimates from Y to X were not really present for either of the LGCM-SR models. However, when fitting the RI-CLPM to the Linear LGCM-SR as the Y slope variance increased so too did the relative bias of the cross-lag parameter estimate. A similar thing was observed for the cross-lag path from X to Y, with a main effect for X slope variance, $\eta_p^2 = 0.5852$, model, $\eta_p^2 = 0.7447$, and the interaction of X slope with model, $\eta_p^2 = 0.7306$. In all cases, the misestimation was biased downwards, thus we can say the RI-CLPM consistently underestimates the cross-lag path when fitting to the Linear LGCM-SR and the underestimation worsened as the slope variance for the predicting variable increased. Figure 12 depicts this effect as given in the table below.

Table 13. Relative Bias in the Cross-Lag Parameter estimates in relation to Predictor's Slope Variance and Model being fit to the Linear LGCM-SR data

Relative Bias in $Y_{t-1} \rightarrow X_t$				
Model	Y Slope Variance	N	Mean	SD
Linear LGCM-SR	5	16298	0.000	0.019
	10	16534	-0.000	0.019
	20	16526	0.000	0.020
Unspecified LGCM-SR	5	17709	0.000	0.019
	10	15150	-0.000	0.020
	20	14132	0.000	0.021
RI-CLPM	5	12569	-0.075	0.013
	10	12148	-0.143	0.030
	20	10701	-0.278	0.055
Relative Bias in $X_{t-1} \rightarrow Y_t$				
	X Slope Variance	N	Mean	SD
Linear LGCM-SR	4	16454	0.000	0.014
	8	16490	-0.000	0.015
	16	16414	-0.000	0.015
Unspecified LGCM-SR	4	17939	0.000	0.014
	8	15783	-0.000	0.015
	16	13269	-0.000	0.016
RI-CLPM	4	15596	-0.108	0.016
	8	10949	-0.171	0.024
	16	8873	-0.275	0.022

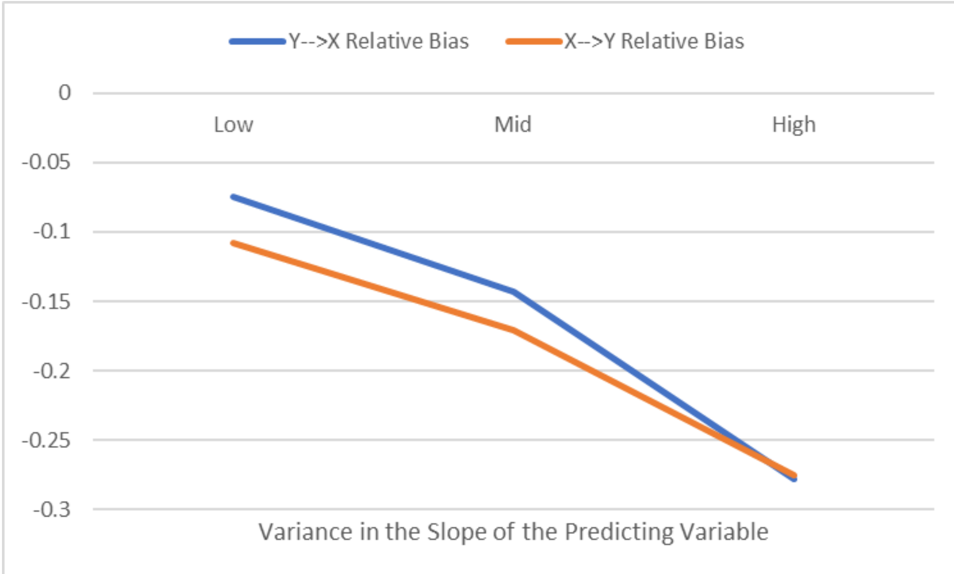


Figure 12. Relative Bias in the Cross-Lag Parameter Estimations for the RI-CLPM when fitting to Linear LGCM-SR data

When fitting to the LGCM-SR-UGT data we again observed that the predicting slope variance influenced the relative bias in our cross-lagged parameters. The main effect of the predicting slope variances, $\eta_p^2 = 0.6471$ and $\eta_p^2 = 0.6072$ for the $Y_{t-1} \rightarrow X_t$ and $X_{t-1} \rightarrow Y_t$ pathways respectively, consistently show an increase in the severity of underestimation as predicting slope variance increased. Corresponding to the increased variance in the predicting slope we also had more issues with inadmissible or non-convergent solutions. More importantly, as reflected in the model main effect, $\eta_p^2 = 0.4136$ and $\eta_p^2 = 0.7305$, along with the model by variance interactions, $\eta_p^2 = 0.6396$ and $\eta_p^2 = 0.6026$, for the $Y_{t-1} \rightarrow X_t$ and $X_{t-1} \rightarrow Y_t$ pathways respectively, we saw that this effect was driven primarily by the RI-CLPM. Namely, no bias was observed when fitting the LGCM-SR-UGT to itself, only when fitting the RI-CLPM. Figure 13 shows the effect of predicting slope variance on relative bias when fitting the RI-CLPM to the LGCM-SR-UGT data.

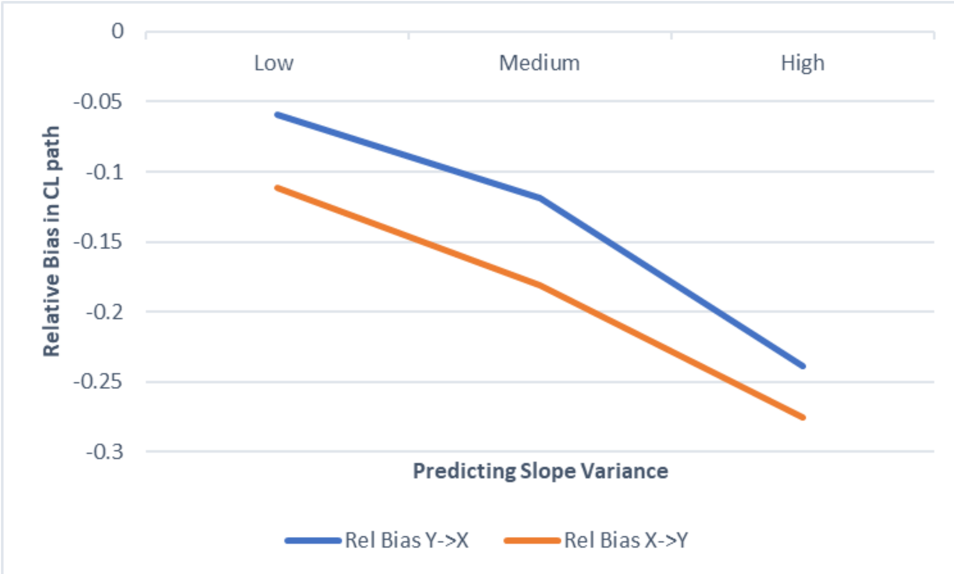


Figure 13. Relative bias in cross-lag pathways across predicting slope variances when fitting the RI-CLPM to LGCM-SR-UGT data

Moreover, when considering the relative bias in estimating the dominated $Y_{t-1} \rightarrow X_t$ pathway, predicting slope variance was moderated by dominance condition, $\eta_p^2 = 0.2436$. This interaction effect was only present when fitting the RI-CLPM to the LGCM-SR-UGT data, $\eta_p^2 = 0.2427$. When a dominant process was present, the fitting of the RI-CLPM to the LGCM-SR-UGT data led to more severe loss of admissible and convergent solutions especially as the Y slope variance increased. Figure 14 depicts the increasing underestimation as a function of Y slope variance and dominance condition when fitting the RI-CLPM to the LGCM-SR-UGT data. As can be seen in the figure below, the severity in underestimation moving across the levels of Y slope variance occurred in both dominance condition but was more severe when no dominant process was present.

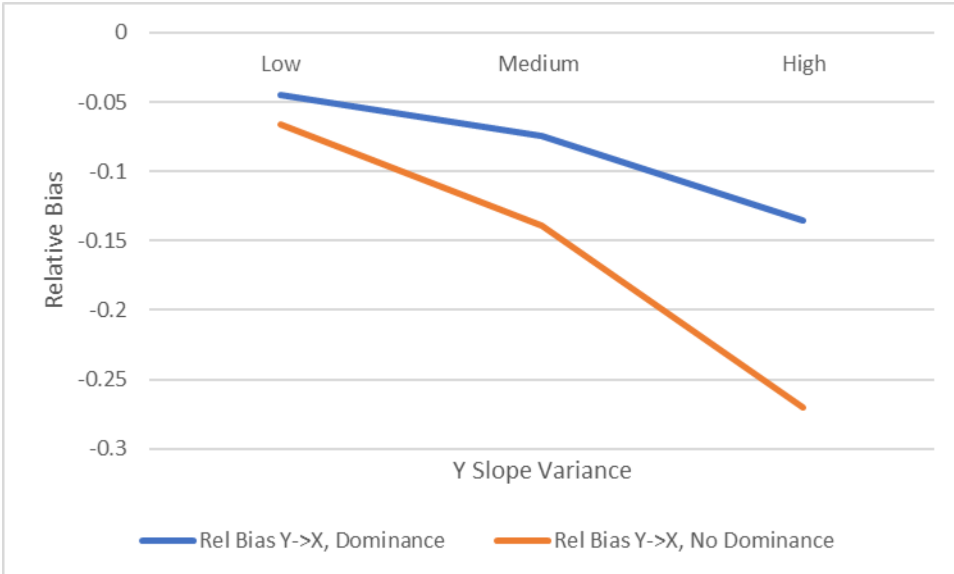


Figure 14. Relative Bias in Y→X pathway by dominance condition and Y slope variance when fitting the RI-CLPM to LGCM-SR-UGT data

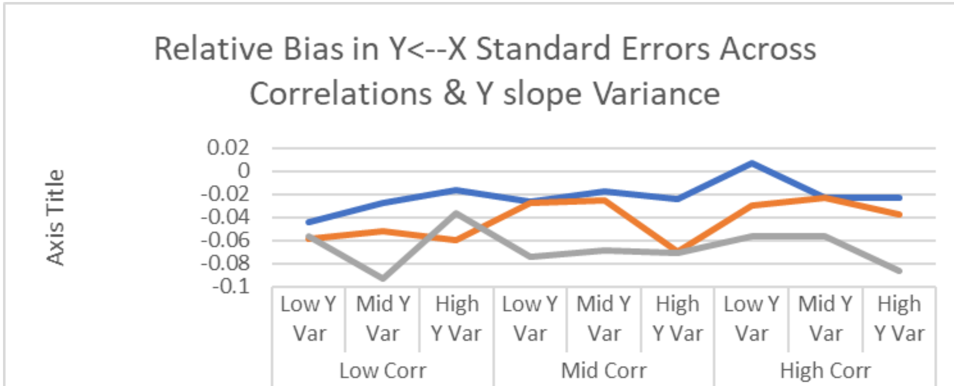
In terms of relative bias in the cross-lag path coefficients, the main story is that when fitting the RI-CLPM to LGCM-SR data we have an underestimation in the cross-lag coefficients, and becomes more severe when the predicting slope variance increases. Further, when fitting specifically to the Unspecified LGCM-SR, the underestimation by RI-CLPM with increases in predicting slope variance is especially pronounced when no dominant process is present.

4.4.3 Relative Bias in the Standard Errors for the Cross-Lag Parameter Estimates

When fitting to the Linear LGCM-SR data, relative bias in the standard errors for both the X to Y and Y to X cross lag estimates had a significant five-way interaction amongst the slope variances and covariance, dominance conditions, and the models, $\eta_p^2 = 0.1376$ and $\eta_p^2 = 0.0948$ respectively. To tease apart this higher order interaction let's begin by exploring the main effects and lower order interactions within each of the respective standard errors' relative biases and then work our way up to the highest order interaction. For the Y to X cross-lag, we first note the significant effect of model, $\eta_p^2 = 0.4962$, wherein the LGCM-SR models were only producing negligible bias (< .005%), while the RI-CLPM was underestimating the Monte Carlo standard deviation by about 10%. This effect likely corresponds to some of the issues we were having with convergence and admissible solutions resulting from a minimization in the errors resulting in negative variance estimates and correlations greater than one. This issue was markedly severe when fitting the RI-CLPM to the Linear LGCM-SR data. It seems that as the error components became overfit the associated standard errors in the cross-lags became underestimated as well.

Moreover, as the predicting variable increased in its slope variance the underestimation became even more severe, $\eta_p^2 = 0.3358$, moving from an underestimation of about 7% for the lowest slope variance to about 15% in the highest variance condition. The LGCM-SR models do not seem to suffer from increases in the predicting slope. The X slope variance by model interaction was moderated by the Y slope variance, $\eta_p^2 = 0.1319$, but the pattern was not clear. At a low X slope variance, the bias became less severe as the Y slope variance increased (from about 8% underestimation in the lowest Y slope variance to about 5% underestimation in the highest Y slope variance). There was no consistent pattern at the other levels of the X slope variance. The LGCM-SR models did not appear to be greatly influenced by the interaction of X and Y slope variances.

The interaction amongst slope variance and models was further moderated by the correlation, $\eta_p^2 = 0.2155$. The exact pattern of this interaction was not entirely consistent. As displayed in Figure 15 below, we can see a few general things; for example, when X slope variance was high and the correlation was low underestimation seemed to become less severe as the Y slope variance increased; to a slighter degree this seemed to occur with low X slope variance. There also appeared to be slight reduction in underestimation as Y slope variance increased for the low and mid-level X slope variance when the correlation was high.

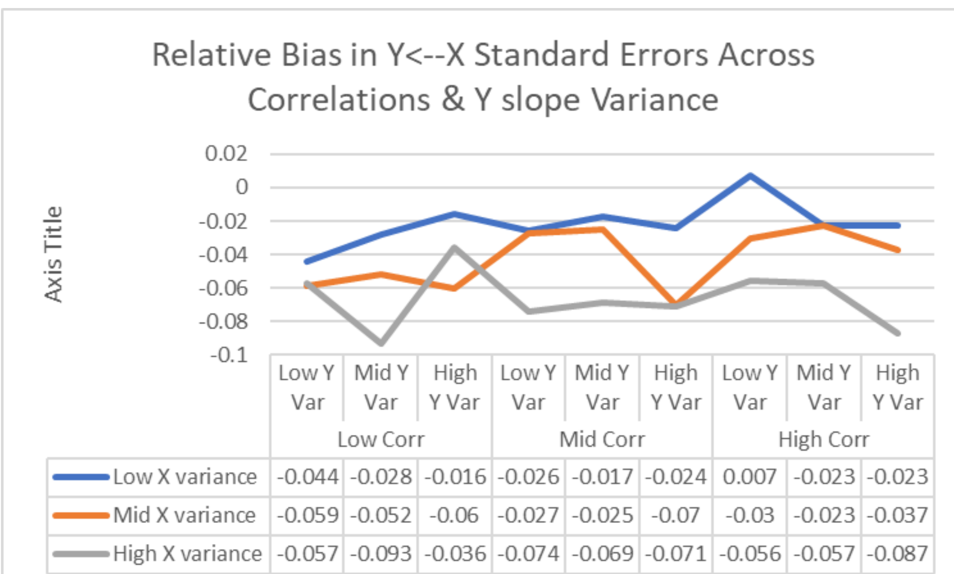


	Low Corr			Mid Corr			High Corr		
	Low Y Var	Mid Y Var	High Y Var	Low Y Var	Mid Y Var	High Y Var	Low Y Var	Mid Y Var	High Y Var
Low X variance	-0.044	-0.028	-0.016	-0.026	-0.017	-0.024	0.007	-0.023	-0.023
Mid X variance	-0.059	-0.052	-0.06	-0.027	-0.025	-0.07	-0.03	-0.023	-0.037
High X variance	-0.057	-0.093	-0.036	-0.074	-0.069	-0.071	-0.056	-0.057	-0.087

Figure 15. Relative Bias in X to Y cross lag standard errors across Y slope variance and X-Y slope correlations at levels of X slope variance when fitting the RI-CLPM to the Linear LGCM-SR data

Having explored up to this four-way interaction with the Linear LGCM-SR data we see that most of the action appeared to be happening when fitting the RI-CLPM. As mentioned before, this four-way interaction was further moderated by the dominance condition, $\eta_p^2 = 0.1376$. The precedence for the dominance condition began with a three-way interaction with X and Y slope variance, $\eta_p^2 = 0.1770$. A closer look at this reveals that higher slope variance conditions within the no dominance condition produced greater underestimation by the standard errors for the X→Y cross-lag path. Notably, when X slope variance was high in the no dominance condition the Monte Carlo standard deviation was underestimated by around 6%-7% across the levels of Y slope variance. When the Y slope variance was high and the X slope variance was at a medium level we found that the X→Y standard errors were underestimating by 5.5%.

This interaction was further moderated by the correlation size, $\eta_p^2 = 0.2640$. Problems with underestimation by the X→Y path standard errors again were only appearing within the no dominance condition. Figure 16 depicts this effect within the no dominance condition but no clear pattern emerges. A few notable changes are that in the low correlation condition we see the underestimation increase in severity moving from low to mid Y slope variance within the high X slope variance condition, and then recover moving from mid to high Y slope variance. In the medium correlation and medium X slope variance condition we found an increase in the underestimation when moving from mid to high Y slope variance.



	Low Corr			Mid Corr			High Corr		
	Low Y Var	Mid Y Var	High Y Var	Low Y Var	Mid Y Var	High Y Var	Low Y Var	Mid Y Var	High Y Var
Low X variance	-0.044	-0.028	-0.016	-0.026	-0.017	-0.024	0.007	-0.023	-0.023
Mid X variance	-0.059	-0.052	-0.06	-0.027	-0.025	-0.07	-0.03	-0.023	-0.037
High X variance	-0.057	-0.093	-0.036	-0.074	-0.069	-0.071	-0.056	-0.057	-0.087

Figure 16. Relative Bias in X to Y cross lag standard errors across Y slope variance and X-Y slope correlations at levels of X slope variance when fitting to the Linear LGCM-SR data in the No Dominance condition

Aside from a couple aberrant cases, it appeared that most of the story with the five-way interaction is occurring when fitting the RI-CLPM to the Linear LGCM-SR. The aberrant cases occurred when fitting the Unspecified LGCM-SR to the Linear LGCM-SR data with no dominance and mid-level X slope variance. When the Y slope variance was 20 and the correlation was at the middle level of 0.3 we had an overestimation of about 11%. Also, when the Y slope variance was at 10 and the correlation was at the high level of 0.5 we had an overestimation of about 5.5%. These may be regarded as flukes, though there may be a more meaningful pattern. Thus, we focus on fitting the RI-CLPM to the Linear LGCM-SR (Table 14 demonstrates these effects).

Most notably we see that in the dominance condition when the X slope variance was at its highest we didn't have observations due to solutions containing negative variance or correlation greater than one in the latent variables. In general, within the dominance condition there was a tendency for misestimation by the X→Y standard errors to become more severe as variances in the slopes increased, but there is no apparent pattern across correlation values. Generally, the tendency was for the bias to be downward, but there was one aberrant case where they overestimated by about 15%. This occurred when the X slope variance was 8, Y slope variance was 20, and the correlation between them was at 0.3.

When moving into the no dominance condition there was far less of an issue with losing observations due to inadmissible solutions as the slope variances increased, thus a direct comparison between the dominance conditions in this context cannot be fully made, because we now have relative bias of the X→Y standard errors in the upper X slope variances. Again, within the no dominance condition, when fitting the RI-CLPM to the Linear LGCM-SR data we had increasing underestimation by the cross-lag path standard errors with increasing slope variances; no clear pattern emerged as we moved across the levels of the correlation between the slopes.

Table 14. Four Way Interaction amongst the slope variances, covariance, and dominance condition for the X→Y path standard error bias when fitting RI-CLPM to the Linear LGCM-SR

Dominance Condition	Correlation of X & Y Slope	Y Slope Variance	X Slope Variance	N	Mean	SD
X dominates Y	0.1	5	4	825	-0.099	0.009
			8	31	0.040	0.012
		10	4	473	-0.055	0.009
			8	7	-0.196	0.011
		20	4	12	-0.188	0.009
	0.3	5	4	977	-0.075	0.009
			8	183	-0.059	0.010
		10	4	924	-0.077	0.009
			8	124	-0.127	0.009
		20	4	410	-0.090	0.009
			8	11	0.149	0.008
	0.5	5	4	998	-0.094	0.009
			8	555	-0.118	0.009
		10	4	998	-0.076	0.009
			8	622	-0.088	0.009
		20	4	979	-0.078	0.009
			8	419	-0.116	0.010
No dominance	0.1	5	4	1000	-0.098	0.010
			8	1000	-0.111	0.010
			16	1000	-0.173	0.009
		10	4	1000	-0.063	0.011
			8	1000	-0.105	0.010
			16	1000	-0.185	0.008
		20	4	1000	-0.046	0.009
			8	997	-0.105	0.009
			16	873	-0.105	0.008
	0.3	5	4	1000	-0.094	0.010
			8	1000	-0.134	0.010
			16	1000	-0.147	0.009
		10	4	1000	-0.070	0.011
			8	1000	-0.099	0.010
			16	1000	-0.158	0.009
		20	4	1000	-0.048	0.010

			8	1000	-0.123	0.009
			16	1000	-0.141	0.008
	0.5	5	4	1000	-0.062	0.011
			8	1000	-0.129	0.011
			16	1000	-0.141	0.009
		10	4	1000	-0.074	0.012
			8	1000	-0.107	0.011
			16	1000	-0.147	0.009
		20	4	1000	-0.028	0.010
			8	1000	-0.086	0.009
			16	1000	-0.143	0.008

The relative bias in the standard errors for the Y→X pathways displayed similar issues as the X→Y paths. Though in this case there was also a much more significant main effect for the dominance condition, the reason for this is likely because the Y→X path was the path which changed values between dominance conditions, i.e. it was the dominated path.

For the Y→X path standard error bias it appeared that the major influences were coming from the different models and the dominance conditions. The main effect for the dominance condition, $\eta_p^2 = 0.4106$, showed that when there is no dominance we underestimated the Y→X Monte Carlo standard deviation by about 7%, whereas in the presence of dominance the bias became negligible. The main effect for the models, $\eta_p^2 = 0.4192$, indicated that when fitting the RI-CLPM to the Linear LGCM-SR data the Y→X standard errors were underestimating by about 14% whereas the LGCM-SR models were not exhibiting misestimation issues. Furthermore, the interaction of the dominance conditions with the models, $\eta_p^2 = 0.3852$, indicated that the RI-CLPM was most severely underestimated in the no dominance condition at about 18%. Table 15 shows these biases, as can be seen, there only appeared to be a major issue with standard error relative bias for Y→X pathway when trying to fit the RI-CLPM to the Linear LGCM-SR data in the no dominance condition.

Table 15. Relative Bias in the Y→X path Standard Errors across the Models and Dominance Conditions

Model	Dominance Condition	N	Mean	SD
Linear LGCM-SR	X dominates Y	26795	0.007	0.032
	No Dominance	22563	0.002	0.023
Unspecified LGCM-SR	X dominates Y	26801	0.007	0.032
	No Dominance	20190	-0.003	0.027
RI-CLPM	X dominates Y	8548	0.031	0.060
	No Dominance	26870	-0.187	0.058

Moreover, the variances and covariance in the slopes moderated the dominance by model interaction ($\eta_p^2 = 0.1183$ for the X slope variance and correlation, $\eta_p^2 = 0.0978$ for the Y slope and correlation, and $\eta_p^2 = 0.0948$ for the X & Y slope variance and correlation). In general, these interactions break out like such: for the RI-CLPM being fit to the Linear LGCM-SR data in the no dominance condition we have increased underestimation as the correlation between the slopes increases, furthermore as the slope variances increase so too does the underestimation. Note the effect on underestimation due to increasing slope variance also occurred within the condition with a dominant process, especially with increases in the Y slope. Furthermore, we also found some underestimation in the upper Y slope variance range for the Unspecified LGCM-SR. Table 16 shows these biases for the Unspecified LGCM-SR and RI-CLPM being fit to the Linear LGCM-SR data.

Table 16. Relative Bias in the Y→X path Standard Errors when fitting the Unspecified LGCM-SR and RI-CLPM to the Linear LGCM-SR data across Slope Variances and Covariance, and Dominance Conditions

Model	Dominance Condition	Slope Correlation	X Slope Variance	Y Slope Variance	N	Mean	SD
Unspecified LGCM-SR	X dominates Y	0.1	4	5	994	0.011	0.020
				10	1000	0.014	0.020
				20	1000	-0.007	0.019
			8	5	992	-0.018	0.020
				10	1000	0.032	0.021
				20	1000	0.019	0.021
			16	5	989	-0.006	0.021
				10	1000	-0.015	0.020
				20	999	0.022	0.021
		0.3	4	5	990	0.028	0.020
				10	1000	0.044	0.021
				20	1000	-0.029	0.019
			8	5	987	0.054	0.021
				10	1000	0.062	0.021
				20	1000	0.012	0.021
			16	5	976	0.008	0.021
				10	1000	-0.016	0.022
				20	1000	-0.034	0.021
		0.5	4	5	958	0.030	0.020
				10	1000	-0.014	0.020
				20	1000	-0.012	0.019
			8	5	958	0.016	0.021
				10	1000	0.020	0.021
				20	1000	-0.023	0.019
			16	5	960	0.017	0.021
				10	998	0.003	0.021
				20	1000	-0.014	0.021
	No Dominance	0.1	4	5	1000	-0.018	0.010
				10	999	0.016	0.010
				20	999	-0.007	0.010
			8	5	1000	-0.029	0.010
				10	758	-0.002	0.010

				20	431	0.034	0.010
			16	5	998	-0.001	0.010
				10	459	-0.017	0.010
				20	163	-0.130	0.008
		0.3	4	5	1000	0.015	0.010
				10	999	-0.005	0.010
				20	1000	-0.003	0.010
			8	5	989	-0.029	0.010
				10	644	-0.028	0.009
				20	561	0.038	0.010
			16	5	981	0.006	0.010
				10	352	-0.029	0.009
				20	94	0.101	0.011
		0.5	4	5	1000	-0.002	0.010
				10	1000	-0.026	0.009
				20	1000	0.006	0.010
			8	5	983	-0.007	0.010
				10	661	0.052	0.010
				20	819	0.007	0.010
			16	5	954	0.010	0.010
				10	280	-0.067	0.009
				20	66	0.096	0.011
3	X dominates Y	0.1	4	5	825	0.044	0.010
				10	473	0.202	0.012
				20	12	-0.143	0.007
			8	5	31	-0.006	0.010
				10	7	0.004	0.006
		0.3	4	5	977	-0.018	0.009
				10	924	-0.001	0.009
				20	410	0.060	0.009
			8	5	183	0.139	0.010
				10	124	0.021	0.009
				20	11	-0.045	0.006
		0.5	4	5	998	0.002	0.009
				10	998	-0.010	0.009

				20	979	-0.011	0.009
			8	5	555	0.117	0.011
				10	622	0.071	0.010
				20	419	0.033	0.009
	No Dominance	0.1	4	5	1000	-0.102	0.008
				10	1000	-0.191	0.009
				20	1000	-0.243	0.008
			8	5	1000	-0.116	0.009
				10	1000	-0.173	0.009
				20	997	-0.239	0.008
			16	5	1000	-0.085	0.009
				10	1000	-0.156	0.009
				20	873	-0.189	0.008
		0.3	4	5	1000	-0.122	0.009
				10	1000	-0.231	0.009
				20	1000	-0.240	0.009
			8	5	1000	-0.113	0.009
				10	1000	-0.214	0.009
				20	1000	-0.251	0.008
			16	5	1000	-0.128	0.009
				10	1000	-0.183	0.008
				20	1000	-0.229	0.008
		0.5	4	5	1000	-0.114	0.009
				10	1000	-0.231	0.010
				20	1000	-0.250	0.009
			8	5	1000	-0.162	0.009
				10	1000	-0.251	0.009
				20	1000	-0.251	0.008
			16	5	1000	-0.111	0.009
				10	1000	-0.222	0.009
				20	1000	-0.263	0.008

In sum, for the Linear LGCM-SR data, the highlight for the X→Y path occurred when fitting the RI-CLPM, the standard errors became increasingly severe in their underestimation as the X slope variance increased; and with the Y→X path, the most severe underestimation occurred when fitting the RI-CLPM to the Linear LGCM-SR with no dominant process.

When fitting to the LGCM-SR-UGT data we again found many higher order interactions, particularly pertaining to which model was being fit to data. Overall, the LGCM-SR-UGT produced negligible relative bias in the standard errors when fit to itself, Mean= -0.001 SD=0.027 for the Y to X path and Mean=-0.003 SD=0.027 for the X to Y path. Thus, the most salient effects occurred in the context of fitting the RI-CLPM to the LGCM-SR-UGT data.

The relative bias for the standard errors of the Y to X pathway exhibited notable underestimation when fitting the RI-CLPM in the non-dominance condition, $\eta_p^2 = 0.3723$. There was a moderation of this effect across the levels of the slope correlation. Specifically, when dominance was present the increasing slope correlations led to a lessening in the overestimation by the standard errors, while the lack of a dominant process exhibited an increase in underestimation as we moved up through slope correlations, $\eta_p^2 = 0.2190$. Figure 17 presents the interaction of dominance and slope correlation when fitting the RI-CLPM to the LGCM-SR-UGT data.

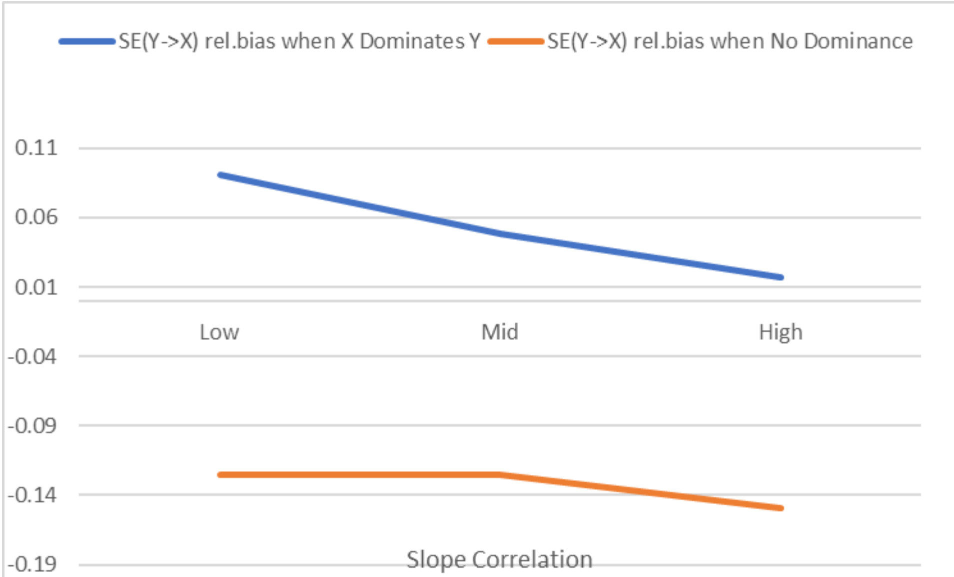


Figure 17. Relative Bias in the Standard Errors for the Y to X cross-lag pathways across the dominance conditions and levels of slope correlation when fitting the RI-CLPM to LGCM-SR-UGT data

In addition, the slope variances had a notable interaction with the dominance condition when fitting the RI-CLPM to the LGCM-SR-UGT data. As the predicting slope increased in variance in the presence of a dominant process the standard errors exhibited a slight decline in overestimation but also yielded fewer and fewer valid solutions. When there was no dominant process, standard error underestimation more severe as the Y slope variance increased, $\eta_p^2 = 0.2408$. This effect was further moderated by increases in the X slope variance, $\eta_p^2 = 0.1628$. This moderation did not have a clear pattern as can be seen in Table 17. In table 17 there is a mixture of underestimation and overestimation. On the whole when X dominates Y we have overestimation, in the highest variance condition we have underestimation, however, the sample size is 4 which doesn't make these results very compelling; in the case where there is no dominance we have underestimation. One consideration worth noting is the difference in the Monte Carlo standard deviation magnitude between the dominance conditions, specifically, in the dominance condition it is much smaller (0.008) than in the non-dominance condition (0.070). There were a few tendencies in the relative bias by the Y to X standard errors that may be noteworthy. For example, we can see that when X slope variance was high we had serious issues in acquiring valid solutions in the presence of a dominant process, in fact no solutions existed in the high X-low Y slope variance cell when a dominant process was present, and only 5 and 4 valid solutions existed at the corresponding medium and high Y slope variance conditions respectively. Given this issue, when X slope variance was high in the dominating condition, we observed the highest bias. The result was apt to be related to the fact that X being the dominating process becomes more poorly estimated as its variance became larger. The patterns within the no dominance condition were a bit clearer, we simply had a situation where underestimation was becoming more severe as the predicting slope variance increased but differed across the levels of the X slope variance, namely at medium X slope variance the severity of underestimation as we moved up the levels of Y slope variance was somewhat mitigated.

Table 17. Relative Bias in the Y to X pathway standard errors across dominance conditions and the slope variances for RI-CLPM fit to LGCM-SR-UGT data

Model	Dominance Condition	X Slope Variance	Y Slope Variance	N	Mean	SD
RI-CLPM	X Dominates Y	4	5	2929	0.032	0.046
			10	2690	0.010	0.044
			20	1776	0.022	0.055
		8	5	1559	0.065	0.052
			10	1409	0.088	0.021
			20	959	0.058	0.056
		16	10	5	0.630	0.018
			20	4	-0.400	0.007
	No Dominance	4	5	3000	-0.092	0.030
			10	3000	-0.139	0.030
			20	3000	-0.191	0.008
		8	5	3000	-0.071	0.026
			10	3000	-0.122	0.021
			20	3000	-0.183	0.029
		16	5	3000	-0.092	0.009
			10	3000	-0.115	0.023
			20	2989	-0.192	0.020

Given that we are dealing with the standard errors, we see that there are many effects pertaining to the variance-covariance components. Though the patterns are not always clear cut, we do find that the combined effect of increasing slope correlation and X slope variance has an additive effect of creating more and more severe underestimation by the standard errors, particularly when fitting the RI-CLPM to the LGCM-SR-UGT data, $\eta_p^2 = 0.2721$, (see Table 18).

Table 18. Relative Bias in the Y to X Standard Errors across the Slope Correlation and X Slope Variance when fitting RI-CLPM to LGCM-SR-UGT data

Model	Slope Correlation	X Slope Variance	N	Mean	SD
RI-CLPM	0.1	4	4700	-0.054	0.118
		8	3230	-0.099	0.065
		16	2989	-0.125	0.051
	0.3	4	5698	-0.071	0.082
		8	4097	-0.043	0.115
		16	3000	-0.125	0.034
	0.5	4	5997	-0.075	0.081
		8	5600	-0.062	0.109
		16	3009	-0.148	0.059

When considering relative bias in the standard errors for the X to Y pathways, we have fewer effects to contend with and again the action is unfolding when fitting the RI-CLPM to the LGCM-SR-UGT data, $\eta_p^2 = 0.3285$. Namely we have interesting interactions between the slope variances, $\eta_p^2 = 0.1082$, as well as an interaction between the X slope variance and slope correlations, $\eta_p^2 = 0.1272$.

Table 19 shows the interaction of the X and Y slope variances when fitting the RI-CLPM to the LGCM-SR-UGT. In general, what appears to be happening with this interaction is that as the X slope variance increases we get more severe underestimation, but, this is offset by increases in the Y slope variance. Figure 18 depicts the influence of slope variances on the bias in the X to Y standard errors when fitting the RI-CLPM to the LGCM-SR-UGT data.

Table 19. Relative Bias in the X to Y standard errors across the levels of Slope Variances when fitting the RI-CLPM to the LGCM-SR-UGT data

Model	Y Slope Variance	X Slope Variance	N	Mean	SD
RI-CLPM	5	4	5929	-0.051	0.026
		8	4559	-0.097	0.021
		16	3000	-0.157	0.010
	10	4	5690	-0.043	0.024
		8	4409	-0.099	0.018
		16	3005	-0.122	0.016
	20	4	4776	-0.033	0.024
		8	3959	-0.074	0.030
		16	2993	-0.113	0.0

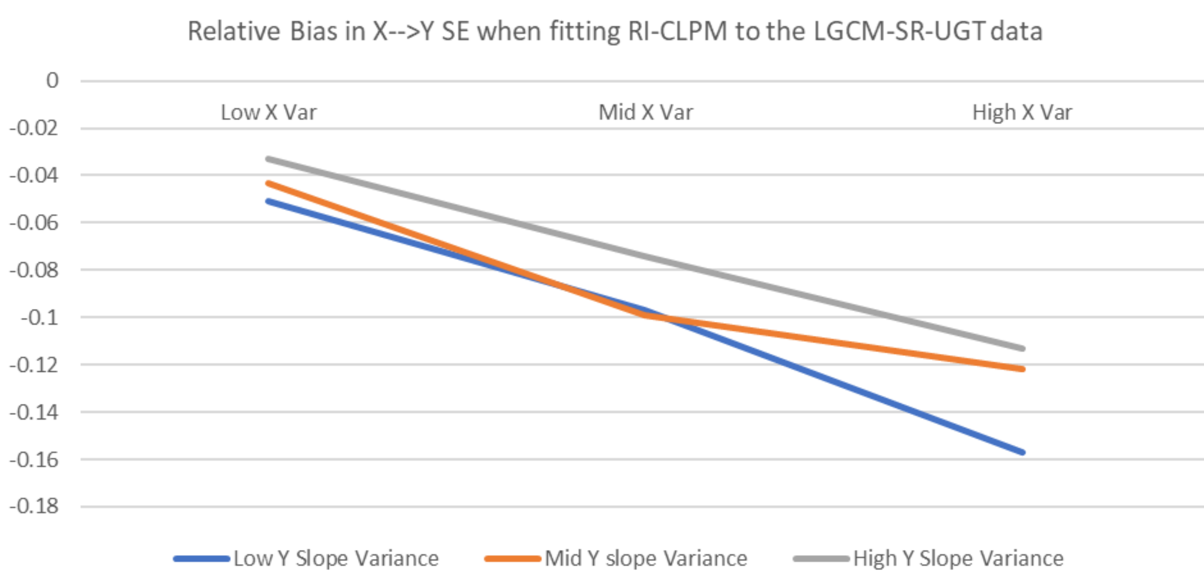


Figure 18. Relative Bias in X-->Y SE when fitting RI-CLPM to the LGCM-SR-UGT data

Table 20 depicts the interaction of the X slope variance with the slope correlation across the dominance conditions. Unfortunately, there does not seem to be any clear pattern; we can only see that for the most part we have a consistent underestimation by the standard errors, and in some combined conditions the underestimation was more severe than it was in other condition combinations. It seems that the most interesting story with relative bias in the X to Y cross-lag standard errors pertained to the interactions amongst variance-covariance components when fitting the RI-CLPM to the LGCM-SR-UGT data.

Table 20. Relative Bias in the X to Y path standard errors across dominance conditions, X slope variances and slope correlations

Dominance Condition	Slope Correlation	X Slope Variance	N	Mean	SD
X Dominates Y	0.1	4	4700	-0.022	0.027
		8	3230	-0.022	0.036
		16	3000	-0.016	0.024
	0.3	4	5698	0.002	0.034
		8	4097	-0.034	0.036
		16	3000	0.002	0.025
	0.5	4	5997	-0.024	0.020
		8	5598	-0.064	0.049
		16	3009	-0.009	0.032
No Dominance	0.1	4	6000	-0.016	0.027
		8	5999	-0.048	0.056
		16	5989	-0.055	0.079
	0.3	4	6000	-0.010	0.037
		8	6000	-0.031	0.051
		16	6000	-0.076	0.063
	0.5	4	6000	-0.048	0.035
		8	6000	-0.039	0.047
		16	6000	-0.071	0.058

In general, the major findings for standard error bias occurs when fitting RI-CLPM to LGCM-SR data. The Monte Carlo standard deviation for the X→Y paths become more severely underestimated as the X slope variance increases and Y→X path standard errors are underestimating when no dominant process is present when fitting the RI-CLPM to the Linear LGCM-SR. Fitting the RI-CLPM to the Unspecified LGCM-SR data also yields underestimation by the standard errors of the Y→X paths when no dominance condition is present, and when there is a dominance condition present these standard errors are overestimating. This overestimation is mitigated as the correlation between slopes strengthens. Thus, when considering the dominated pathway, there is some indication here that we have a higher risk for type 1 error when no dominance is present, and risk type 2 error when dominance is present. The type 2 risk is offset when two processes are more closely related in the dominated condition while the type 1 risk in the non-dominated condition becomes worse as the correlation between the slopes increases. The X→Y path standard errors demonstrate increased underestimation as the X slope variance increases; however, this underestimation is counteracted when the Y slope variance increases.

4.5 A NOTE CONCERNING THE RI-CLPM GENERATED MODELS

As we noted before, the LGCM-SR models could never derive admissible or convergent solutions when being fit to the RI-CLPM data. The likely reason for this is that we have an overfitting problem. The RI-CLPM had no issues in fitting to itself, as implied by always estimating a valid solution across replications and conditions. Only two outcomes presented any significant effect: Relative Bias in the X to Y standard errors and the information criteria. Figure 19 shows that we get slightly worse misestimation by the standard errors for the X to Y path when a dominant process was presented, but it should also be noted that the variance in this bias encompasses the distribution of the bias when there is no dominant process, $\eta_p^2 = 0.1282$.

Additionally, as addressed before, BIC had a better fit when a dominant process was present.

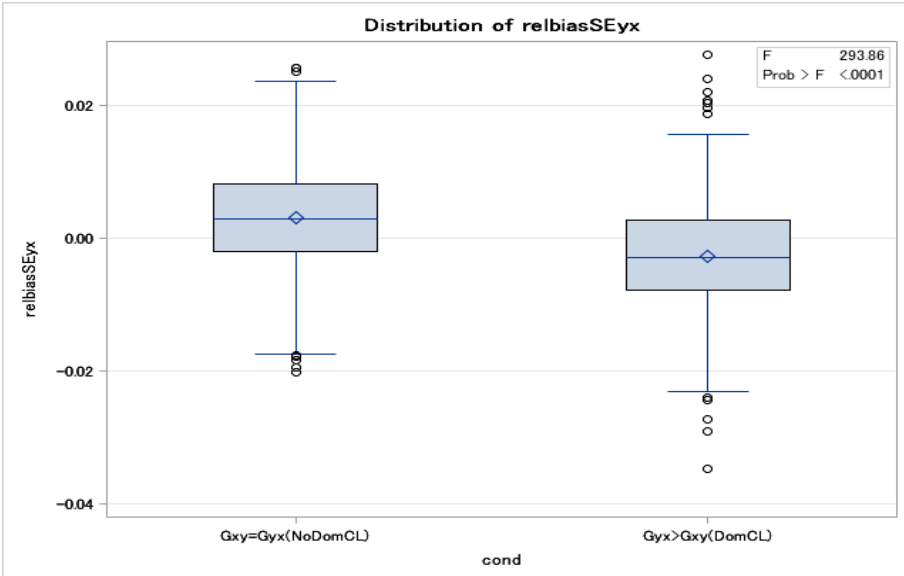


Figure 19. Relative Bias in the X to Y path standard error across dominance conditions when fitting the RI-CLPM to itself

5.0 DISCUSSION

In this study, we set out to explore the performance of models that disaggregate between- and within-person effects while controlling for autoregression and utilizing cross-lag components to assess causal dominance amongst developmental processes. More specifically, our interest was in how variation in the between and within person components affected the fit of the model and the conclusions we derive based on the cross-lag components.

The models under consideration each account for different aspects of capturing the between person effects with the purpose of making correct inference about the within person effects in terms of how two developmental processes causally relate to one another. The Random Intercept Cross-Lag Panel Model (RI-CLPM) accounts for trait stability at the between person level, and captures the form of growth over time by fitting time specific means, but this model does not account for between-person variability in change over time. The Latent Growth Curve Models with Structured Residuals (LGCM-SR) account for between person change and variability in change by fitting a latent slope. With a linear specification, we restrict the type of change to be of a specific form, while we can use an unspecified trajectory to more accurately capture the functional form of growth amongst individuals. In this way, the RI-CLPM and Unspecified LGCM-SR share a commonality by capturing a non-linear form of growth, whereas the Linear and Unspecified LGCM-SR share a commonality by capturing variability in change between individuals. In evaluating the relative performance amongst the models, we generated data sets from each of these models and analyzed those datasets with each of the datasets in turn.

5.1 SUMMARY OF MAJOR FINDINGS

From our applied example with the LSAY data, we saw that disentangling between and within-person effects using RI-CLPM and LGCM-SR improved the model fit over CLPM; however, when considering the FI-CLPM this was only the case by SRMR criteria. This may be the case because FI-CLPM is fitting more covariances and SRMR is considering the absolute differences in model and observed covariances with no penalty for parameters. Moreover, fitting slopes led to even better fit, especially when fitting the unspecified trajectory.

Hamaker et al. (2015) and Berry & Willoughby (2016) conducted studies demonstrating that the inclusion of between-person components as is done with the RI-CLPM and LGCM-SR, respectively, could lead us to different results based on the cross-lag parameter estimates derived from a standard cross-lag model (CLPM). More specifically, we could arrive at opposite conclusions about how processes relate to one another depending on whether we do or do not include components that explicitly account for between-person variance. The findings from our study reinforce the results from these prior studies. When we did not disentangle the effects, we concluded that self-concept was dominating task value, but when we did disentangle the effects we conclude that task value is dominating self-concept. In the simulation study, we don't observe this switching because we were only considering models that do disentangle between and within person effects. Thus, we did not generate any further indication that additional accounting of between person variance bears upon the conclusions we derive about the cross-lag paths.

Additionally, when we fit the intercepts we see a reduction in our autoregressive parameters, and when we include a slope we see a reduction in the cross-lag parameters. This implies that the between person variance and covariance components have a considerable influence on the within-person autoregressive and cross-lag components, ultimately leading us to arrive at different conclusions about the nature of the processes under study.

Perhaps one of the most salient results from the simulation study pertains to when convergent and admissible solutions could be derived. Curran (2003) drew the parallel between multilevel models and latent variable models that place a conceptual underpinning for the approach taken in the models discussed here, wherein latent variables are being used to fit random effects. The general spirit of both latent variables and random effects is to absorb variance; however, this variance is estimated out of the residuals. Because these components are estimates, when we have very low or non-existent error remaining in the residual variance due to a source of error we can end up with negative variance estimates. The phenomena of negative variance estimates are well known within multilevel modeling (Goldstein, 2005). Similarly, we may postulate that when a covariance component is fit it may be attempting to estimate a component relating error structures that are relatively low thus leading to a correlation value exceeding an absolute value of one. In our study, inadmissible solutions were due to negative variance estimates in the exogenous latent variables and correlations amongst latent variables exceeding an absolute value of 1. These issues reflect mis-fitting problems such as overfitting complex models to simpler data and misallocating between- and within-person variance when down-fitting simpler models to more complex data in a latent variable/multilevel context. The general problem is based on the issue of fitting more latent/random effect parameters than the actual number of observed components, which exacerbates the problem of pulling out more variance than is present. These issues are especially pronounced when no dominant process is present, and very often the negative variance issues occur with the X slope variance. Perhaps this is because the X slope has less variance than the Y slope, thus more readily overfit.

The specifics of acquiring admissible and convergent solutions varies across generating models. When fitting to the Linear LGCM-SR data we have information from each analytic model, with the Unspecified LGCM-SR yielding valid solutions 87% of the time and the RI-CLPM only yielding valid solutions 65% of the time. When fitting the RI-CLPM, acquiring valid solutions worsens as the slope variance and covariance increase. The Unspecified LGCM-SR generating model does well at fitting to itself, but the Linear LGCM-SR never fits. However, the RI-CLPM fits the Unspecified LGCM-SR data 71% of the time. This gives some indication that capturing the functional form may be more important than capturing variability in the trajectory. In consideration of overfitting, there is the finding that neither of the LGCM-SR models can be validly fit to the RI-CLPM data. This is best explained by bearing in mind that the RI-CLPM has no variability in change, so when we try to explain such variability we are forcing a situation of overfitting.

When considering Model Fit Indices, a general finding is that when fitting to LGCM-SR data, all model fit indices showed high consistency in preferring an LGCM-SR model to an RI-CLPM model. Since we were only able to fit an Unspecified LGCM-SR to the Linear LGCM-SR data and not the other way around, we can only consider differences in correct selection between fitting Linear vs. Unspecified LGCM-SR to Linear LGCM-SR data. In this case, we find that BIC does best at making the correct selection, which is likely due to the more strenuous penalty for model complexity exhibited by BIC. While AIC also does well at correctly choosing the Linear over the Unspecified model, it is second best to BIC. On the other hand, the penalties applied by both AIC and BIC also make it a tad less desirable for correctly selecting over the RI-CLPM.

The relative fit indices are perfect at correctly choosing the Linear over the RI-CLPM, but they perform fair at best in correctly choosing over the Unspecified LGCM-SR. The main reason for this is that often both Unspecified and Linear analytic models indicated at or near perfect fit, and we tend to only make the correct selection due to parsimony when both are exactly perfect. When choosing the Unspecified over the Linear, this is usually done when the index values are only off by extremely small decimal place values but still favoring the Unspecified model. A similar thing happens when using the RMSEA criteria, except in this case it is values within decimals of 0. TLI is also considerably poorer at correct model selection than CFI. This may be due to the fact the CFI will be higher than TLI because it doesn’t penalize for complexity as TLI does. Thus, more often we have perfect CFI values for both LGCM-SR models leading us to correctly choose Linear over Unspecified for parsimony. An important consideration to be noted when considering the model selection is that the non-linear trajectory did not exhibit extreme deviation from linearity thus fit indices, in general, may not do well at distinguishing the Linear from the Non-Linear LGCM-SR presented here.

Relatedly, when perfect fit is not exhibited by both LGCM-SR models, more often we expect the greater variance explained by the Unspecified model to make it erroneously chosen over Linear. SRMR is the worst criteria for correctly choosing the Linear over the Unspecified model. This is likely because the SRMR is more concerned with how much variance is explained regardless of model complexity. As we know, the Unspecified model will tend to match the observed model covariance matrix by explaining out any extra sample based variance that occurs within the slope by allowing for the overfitting of variance in the trajectories.

The prior literature (Hamaker et al., 2015; Berry & Willoughby, 2016; Curran et al., 2013) only considered the effect of not disaggregating within and between-person variance on the estimation of cross-lag components. In this study, we moved beyond disentangling between-person effects from within-person effects in general and made a further exploration of the effects of accounting for between-person differences in change in addition to between-person stability. The characteristics of our simulation study were founded on the distinctions between the RI-CLPM (Hamaker et al., 2015) and the LGCM-SR (Curran et al., 2013). Moreover, we made a further distinction of the trajectory we fit to the LGCM-SR. Curran et al. (2013) presented only on the linear model, while in our study we also accounted for the possibility of non-linear trajectory. The admission of this non-linear trajectory is an important consideration, because as mentioned before, the RI-CLPM (Hamaker et al., 2015) fit time specific means to capture the trajectory, yet this model does not capture variation around the non-linear trajectory as a LGCM-SR with an unspecified growth curve would.

Our findings show some consistency with this prior research in so far as we demonstrated that within-person cross-lag components can become misrepresented when we fail to accurately capture the nature of between-person differences in stability and change over time. Despite the limitations of our results due to non-convergent and inadmissible solutions we have some indication that accounting for between-person differences in change is important given that RI-CLPM could always be fit to LGCM-SR data, while LGCM-SR models could never be fit to RI-CLPM data; unfortunately, we cannot address at any depth whether fitting the non-linear trajectory via RI-CLPM is better than fitting it with the Linear LGCM-SR, since only RI-CLPM was estimated when fit to the non-linear LGCM-SR.

In the following the results of the cross-lag parameter outcomes from our simulation are presented. The conclusions we derive from our cross-lag components are assessed in three different ways: a factor of dominance, relative bias in the path coefficients, and relative bias in the standard errors of the cross-lag path coefficients. Similar to findings from the simulation study comparing RI-CLPM to CLPM (Hamaker et al., 2015) we see that RI-CLPM tends to underestimate path coefficients when fit to the LGCM-SR and the extent of this underestimation has a notable relation to the dominance condition and between-person variance components. In the case of this study, the manipulated between-person component is in the slope terms, whereas in their simulation it was the intercept. When fitting the RI-CLPM to the Unspecified LGCM-SR data we have notable underestimation of the factor of dominance especially in the presence of a dominant process. Interestingly, the underestimation becomes more severe as the X slope variance increases (i.e., dominant process variability), but becomes less severe as the Y slope variance increases (i.e., dominated process variability).

In terms of relative bias in the cross-lag path coefficients, the main story is that when fitting the RI-CLPM to LGCM-SR data we have an underestimation in the cross-lag coefficients, and becomes more severe when the predicting slope variance increases. Further, when fitting specifically to the Unspecified LGCM-SR, the underestimation by RI-CLPM with increases in predicting slope variance is especially pronounced when no dominant process is present. The relative bias in the standard errors for the cross-lag pathways showed many interactions and effects and in many cases the patterns were hard to discern. As with our other outcomes most of the action occurs when fitting RI-CLPM to LGCM-SR data. The Monte Carlo standard deviation for the X→Y paths become more severely underestimated as the X slope variance increases and Y→X path standard errors are underestimating when no dominant process is present when fitting the RI-CLPM to the Linear LGCM-SR. Fitting the RI-CLPM to the Unspecified LGCM-SR data also yields underestimation by the standard errors of the Y→X paths when no dominance condition is present, and when there is a dominance condition present these standard errors are overestimating. This overestimation is mitigated as the correlation between slopes strengthens. Thus, when considering the dominated pathway, there is some indication here that we have a higher risk for type 1 error when no dominance is present, and risk type 2 error when dominance is present. The type 2 risk is offset when two processes are more closely related in the dominated condition while the type 1 risk in the non-dominated condition becomes worse as the correlation between the slopes increases. As regards the type 1 risk, it is worth noting that both the standard errors and path coefficients are being underestimated, hence the type 1 risk may not be of too much concern. The X→Y path standard errors demonstrate increased underestimation as the X slope variance increases; however, this underestimation is counteracted when the Y slope variance increases.

The most common issue in the cross-lag path estimates pertains to underestimation when fitting the RI-CLPM to LGCM-SR data. The underestimation issues are most pronounced in non-dominance conditions and worsen as slope variances increase. A particularly salient point pertaining to the cross-lags is that we have a greater risk of type 1 error for our paths when no dominance is present which becomes worse as variance and covariance amongst slopes increases; we have a greater type 2 risk for a dominated path that is offset by increasing strength in the relation between two processes. Convergence rates revealed to us that when fitting to the Unspecified LGCM-SR data, that only RI-CLPM can be fit. Hence, in terms of the relative performance of Linear LGCM-SR to RI-CLPM when fitting to Unspecified LGCM-SR data, it would seem appropriate to say that capturing the functional form of the trajectory as is done by fitting means in the RI-CLPM is more important than capturing the variability in the trajectory as is done with the Linear LGCM-SR.

5.2 CONSIDERATIONS, LIMITATIONS, AND FUTURE RESEARCH

One of the most salient issues presented by this study relates to the tendency towards non-convergent and inadmissible solutions. As noted, the problems appear to be based on overfitting complex models to simpler data and misallocating between- and within-person error when down-fitting simpler models to more complex data. Overfitting, in general, is an issue that needs to be considered when latent variables are used to model data, while down-fitting need to be considered when fitting multilevel models. Another consideration that may bear upon the inadmissible solutions pertains to the relative size of the different variance components in relation the total variance in the outcomes. In our models the residuals have considerably larger share of the total variance than the intercept and slope variances. Future research would benefit by considering the influence of relative sizes of variance components on the admissibility of solutions. An additional finding was that trying to fit a linear slope to a non-linear trajectory led to severe problems with estimation, such that 100% of the solutions were inadmissible. Thus, careful consideration to the form of growth and the variability in growth is very important.

These mis-fitting problems negatively impacted other issues for understanding the results of this study. One key issue is that there is a potential confounding of model fit and cross-lag bias with the varying sample sizes across conditions and models that are due to invalid solutions. Further research will need to more explicitly explore the influence of sample size on the information we derive from fitting these models to data. Because of the prevalence of issues in underestimation, it would be of interest to explore the extent to which this underestimation is associated with our model mis-fitting issues. For practical reasons, it is important to explore ways to prevent these mis-fitting problems, and to explore how to set up our estimation method in such a way that these misfit problems are mitigated. This prompts further research that can tease apart the exact factors leading to the mis-fitting of these models. With such an understanding, we could establish better generating conditions to improve the amount of valid solutions which would allow us to have a better comparison and examination of the models under consideration in this study.

A recurring finding throughout this study was that the dominance condition was a very important factor influencing not only model fit, but also bias in our cross-lag components as well as the sheer amount of valid solutions that we could estimate. Further exploration of what it is about these dominance conditions that bears upon the results from our models would be helpful. Moreover, it would be of interest to examine how changing the factor of dominance influences the performance of these models. As a final consideration, it may be of interest to conduct future studies to explore more aspects of these models. For example, how might the sign of our correlations and coefficients influence the relative performance of these models. Future research that makes closer consideration of not just the between person change but also the between person stability would be an important development in our understanding of the models under consideration. It is plausible that in this study the autoregressive parameters being set to 1.2 may have been a bit large. Thus, it would be important to closely examine how changes in the autoregressive components bear upon the performance of these models. Another limitation pertains to the fact that we kept so many parameters constant across time, when in reality phenomena tends to be much more complicated than this.

This study serves as a beginning step in exploring the nature and performance of these recent developments in multilevel autoregressive cross-lag modeling. In time, more and more issues will require exploration. The findings presented here are intended to give some insight into how we can apply these models and understanding the results we derive from such modeling. Future advances in this type of modeling will be needed to expand the capacity and potential for exploring our research questions that are served well by such models. In time, it will be exciting to see and contribute to the developments in multilevel autoregressive cross-lag modeling from both a fixed and random effects perspective.

5.3 MODEL FITTING RECOMMENDATIONS FOR APPLIED RESEARCHERS

In the following we overview some consequences of this research for informing applied research. To start, of central concern, is to get a sense of which model the data is most closely aligned with. The recommendation for this is to fit different components of the models to data in sequence. By fitting the growth curves to each process independently we can gauge both the shape and variance in the trajectories, which is exactly the feature that the models evaluated differ on.

The first thing to observe would be the shape of the trajectory for each process. To determine whether we have a linear or non-linear trajectory our findings suggested that using BIC in model selection is preferable, thus one would fit both the LGCM-UGT and LGCM-LIN to each process and compare the respective BIC values, opting for the model yielding the lower BIC value. As our findings suggest, if we have a non-linear trajectory then we should fit either a Random Intercept (RI) with estimated means or LGCM-UGT because we found that fitting a

LGCM-LIN never yielded admissible solutions. If we have a linear trajectory then any model can be fit, however, fitting the LGCM-UGT would not be parsimonious so we would be choosing between LGCM-LIN or RI with estimated means. Upon determining a trajectory shape for each process, the next step would be to determine whether slope variance will need to be accounted for. Recall, increasing slope variance, when not accounted for, is linked to increasing underestimation in the cross-lag components, thus it is important that such variance is modeled. Software for fitting Structural Equation Models yield output that allows for the testing of the hypothesis that slope variance is equal to zero. Further, given that we are considering multilevel models we can consider whether including the slope variance is appropriate via intra-class correlation, likelihood ratio test, or BIC. If we find that our slope variance was notable, then given a non-linear trajectory we would fit LGCM-UGT to the process and LGCM-LIN given a linear trajectory. If slope variance was non-significant, then given either linear or non-linear trajectory we would fit the RI with estimated means. Alternatively, we could conceivably fit the growth curves (LGCM) with slope variance constrained to zero, however, this was not explored in this study.

Fitting the proper between-person model to each process is the key to estimating the CLPM model as we desire (namely, the effects of intra-individual changes in one process on intra-individual changes in the other). Thus once we have found the proper between-person model for each process individually, the next recommend step would be to bring the two between-person models together. This allows us to estimate the between-person variance-covariance components. The importance of this step is to address issues we encountered wherein between-person error was potentially being misallocated to the within-person components. Once these components are fit we could fix those between-person variance components and then structure our residuals for estimation of the within-person model (CLPM). The figure below (Figure 20) depicts a flow chart that graphically summarizes the recommended steps in selecting and fitting models.

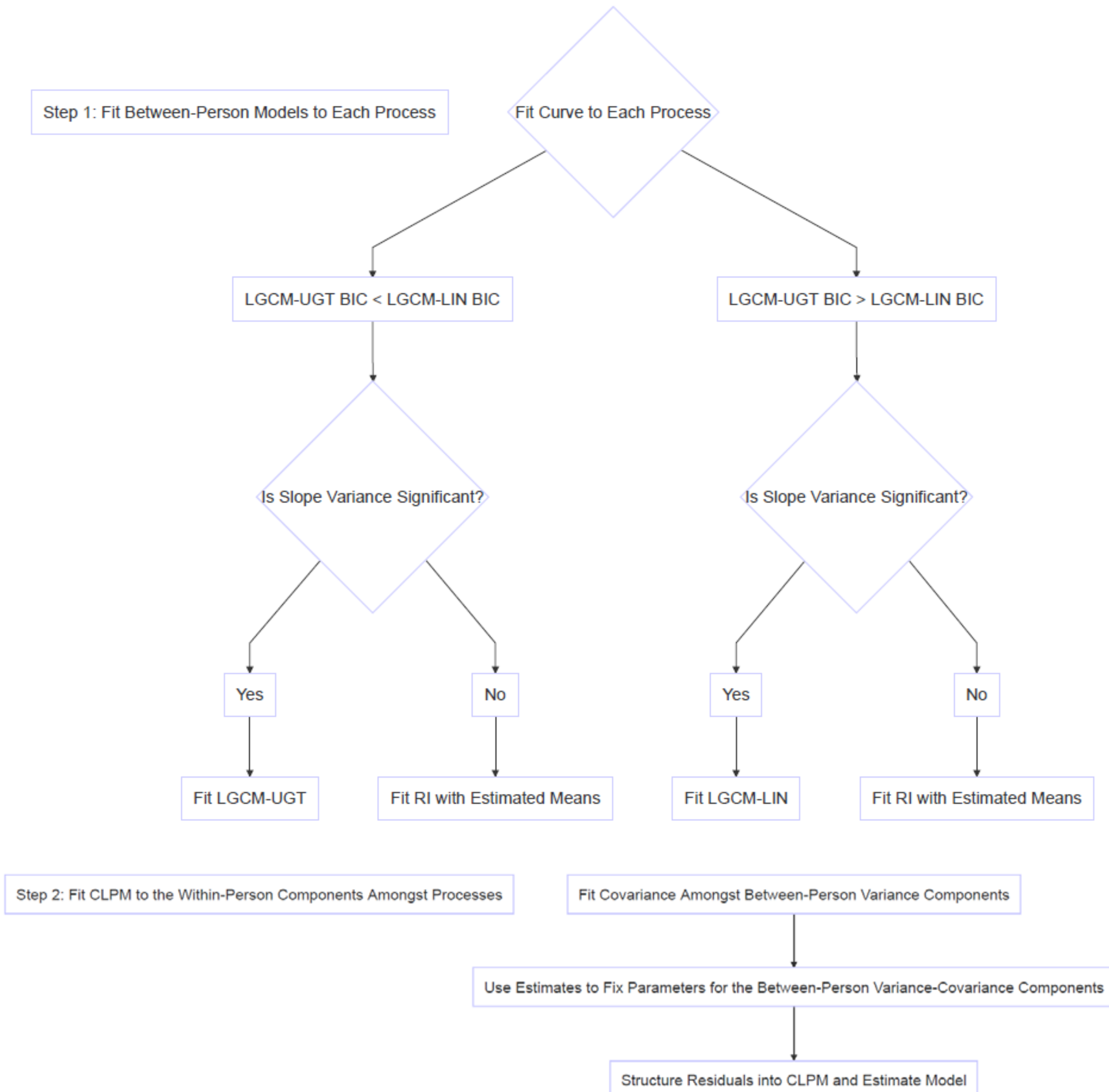


Figure 20. Flowchart for Selecting and Fitting Models

BIBLIOGRAPHY

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle (1973). In *Selected Papers of Hirotugu Akaike* (pp. 199-213). Springer New York.

Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS*. SAS Institute.

Allison, P. (2015). “Don’t put Lagged Dependent Variables in Mixed Models”. http://statisticalhorizons.com/lagged-dependent-variables

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum Likelihood for Cross-lagged Panel Models with Fixed Effects. *Socius*, *3*, 2378023117710578.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Bentler, P. M., & Speckart, G. (1981). Attitudes" cause" behaviors: A structural equation analysis. *Journal of Personality and Social Psychology*,*40*(2), 226.

Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later?. *Journal of Educational Psychology*, *96*(1), 97.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research*, *32*(3), 336-383.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). John Wiley & Sons.

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

Burke, J. D., Pardini, D. A., & Loeber, R. (2008). Reciprocal relationships between parenting behavior and disruptive psychopathology from childhood through adolescence. *Journal of abnormal child psychology*, *36*(5), 679-692.

Cacioppo, J. T., Hawkley, L. C., & Thisted, R. A. (2010). Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychology and aging*, *25*(2), 453.

Campbell, D. T. (1963). From description to experimentation: Interpreting trends as quasi-experiments.

Cohen, Jacob. (1988). Statistical Power analysis for the Behavioral Sciences, 2nd ed. Lawrence Erlbaum Associates

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of abnormal psychology*, *112*(4), 558.

Collins, L. M., & Sayer, A. G. (2001). *New methods for the analysis of change*. American Psychological Association.

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models

Kenny, D. A., & Zautra, A. (2001). Trait–state models for longitudinal data.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data.

Curran, P. J. (2003). Have multilevel models been structural equation models all along?. *Multivariate Behavioral Research*, *38*(4), 529-569.

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of consulting and clinical psychology*, *82*(5), 879.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual review of psychology*, *62*, 583-619.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121-136.

Eccles, J. S., Lord, S. E., Roeser, R. W., & Barber, B. L. (1997). “The association of school transitions in early adolescence with developmental trajectories through high school.” In *Health Risks and Developmental Transitions During Adolescence.* Eds. J.Schulenberg, J.I. Maggs, and K. Hurrelmann. (New York: Cambridge University Press), 283-321.

Goldstein, H. (2005). *Multilevel models*. John Wiley & Sons, Ltd.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological methods*, *20*(1), 102.

Hedges, L. V. (2007). "Effect sizes in cluster-randomized designs." *Journal of Educational and Behavioral Statistics*, 32(4): 341-370.

Kenny, D. A. (1973). Cross-lagged and synchronous common factors in panel data. Structural equation models in the social sciences. New York: Seminar Press.

Luhmann, M., Schimmack, U., & Eid, M. (2011). Stability and variability in the relationship between subjective well-being and income. *Journal of Research in Personality*, *45*(2), 186-197.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, *1*(2), 130.

Mackinnon, S. P. (2012). Perceived social support and academic achievement: Cross-lagged panel and bivariate growth curve analyses. *Journal of youth and adolescence*, *41*(4), 474-485.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological methods*, *12*(1), 23.

McArdle, J. J. (1989). A structural modeling experiment with multiple growth functions. In *Abilities, motivation, and methodology: The Minnesota symposium on learning and individual differences* (pp. 71-117). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Miller, Jon D. Longitudinal Study of American Youth, 1987-1994, 2007-2011. ICPSR30263-v6. Ann Arbor, MI: Inter-university Consortium for Political and Social Research

Muthén, L. K., & Muthén, B. O. (1998-2015). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén

Nelemans, S. A., Hale III, W. W., Branje, S. J., Hawk, S. T., & Meeus, W. H. (2014). Maternal criticism and adolescent depressive and generalized anxiety disorder symptoms: a 6-year longitudinal community study. *Journal of abnormal child psychology*, *42*(5), 755-766.

Pardini, D. A., Fite, P. J., & Burke, J. D. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and African-American ethnicity. *Journal of abnormal child psychology*, *36*(5), 647-662.

Petersen, I. T., Bates, J. E., & Staples, A. D. (2015). The role of language ability and self-regulation in the development of inattentive–hyperactive behavior problems. *Development and psychopathology*, *27*(01), 221-237.

Preacher, K. J. (2008). *Latent growth curve modeling* (No. 157). Sage.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*(2), 245.

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological methods*, *21*(2), 206.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2-3), 144-164.

Shaffer, A., Lindhiem, O., Kolko, D. J., & Trentacosta, C. J. (2013). Bidirectional relations between parenting practices and child externalizing behavior: A cross-lagged panel analysis in the context of a psychosocial treatment and 3-year follow-up. *Journal of abnormal child psychology*, *41*(2), 199-210.

Steiger, J. H., & Lind, J. C. (1980, May). Statistically based tests for the number of common factors. In *annual meeting of the Psychometric Society, Iowa City, IA* (Vol. 758, pp. 424-453).

Usami, S., Hayes, T., & McArdle, J. J. (2015). On the Mathematical Relationship Between Latent Change Score and Autoregressive Cross-Lagged Factor Approaches: Cautions for Inferring Causal Relationship Between Variables. *Multivariate behavioral research*, *50*(6), 676-687

Usami, S., Hayes, T., & McArdle, J. J. (2016). Inferring longitudinal relationships between variables: Model selection between the latent change score and autoregressive cross-lagged factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(3), 331-342.

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological methods*, *20*(1), 63.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary educational psychology*, *25*(1), 68-81.

Williams, R., Allison, P., & Moral-Benito, E. (2015). ” Linear Dynamic Panel Model Estimation Using Maximum Likelihood and Structural Equation Modeling”. Stata Conference. July 30, 2015.