ijbd

International Journal of Behavioral Development
2021, Vol. 45(5) 440–452
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/01650254211022862
journals.sagepub.com/home/jbd
SAGE

Study length, change process separability, parameter estimation, and model evaluation in hybrid autoregressive-latent growth structural equation models for longitudinal data

D. Angus Clark[1], Amy K. Nuttall[2], and Ryan P. Bowles[2]

Abstract

Hybrid autoregressive-latent growth structural equation models for longitudinal data represent a synthesis of the autoregressive and latent growth modeling frameworks. Although these models are conceptually powerful, in practice they may struggle to separate autoregressive and growth-related processes during estimation. This confounding of change processes may, in turn, increase the risk of the models producing deceptively compelling results (i.e., models that fit excellently by conventional standards despite highly biased parameter estimates). Including additional time points provides models with more raw information about change, which could help improve process separability and the accuracy of parameter estimates to a degree. This study thus used Monte Carlo simulation methods to examine associations between change process separability, the number of time points in a model, and the consequences of misspecification, across three prominent hybrid autoregressive-latent growth models: the Latent Change Score model (LCS), the Autoregressive Latent Trajectory Model (ALT), and the Latent Growth Model with Structured Residuals (LGM-SR). Results showed that including more time points increased process separability and robustness to misspecification in the LCS and ALT, but typically not at a rate that would be practically feasible for most developmental researchers. Alternatively, regardless of how many time points were in the model process separability was high in the LGM-SR, as was robustness to misspecification. Overall, results suggest that the LGM-SR is the most effective of the three hybrid autoregressive-latent growth models considered here.

Keywords

Latent change score model, autoregressive latent trajectory model, latent growth model with structured residuals, asymptotic covariance, bias

Introduction

Hybrid autoregressive-latent growth longitudinal structural equation models (SEMs) (hereafter referred to as "hybrid models" for simplicity)—such as the Latent Change Score model (LCS; McArdle, 2001), Autoregressive Latent Trajectory Model (ALT; Bollen & Curran, 2006), and Latent Growth Model with Structured Residuals (LGM-SR; Curran et al., 2014)—synthesize the autoregressive and latent growth modeling traditions, capturing both the extent to which past status relates to future status and absolute change over time (Bollen & Curran, 2004). Hybrid models can struggle to differentiate between autoregressive and growth processes during estimation, however, which may explain why some hybrid models produce extremely biased, yet well fitting, solutions in the face of even minor, routine misspecifications (Clark et al., 2018; Jongerling & Hamaker, 2011; Voelkle, 2008). There is preliminary evidence though that hybrid models are better at separating change processes as additional time points are included (Clark et al., 2018). The present study thus uses Monte Carlo simulation methods to examine the interplay between change process separability, number of time points, parameter bias, and model fit, across three popular hybrid models (LCS, ALT, and LGM-SR).

Hybrid Autoregressive-Latent Growth Longitudinal SEMs

The most commonly used SEM for analyzing longitudinal data in the social and life sciences are the autoregressive and latent growth model (Bollen & Curran, 2004). In autoregressive models (AR), scores at one time point are regressed on scores at one or more previous time points. The regression coefficients capture the extent to which past status is related to future status or the extent to which rank-order stability is preserved across time (Biesanz, 2012). In latent growth models (LGM), scores across time are used as indicators of two latent factors, an intercept and a slope factor (Bollen & Curran, 2006). These latent factors are conceptually distinct from

[1] University of Michigan, USA
[2] Michigan State University, USA

Corresponding author:
D. Angus Clark, Department of Psychiatry, Addiction Center, University of Michigan, Rachel Upjohn Building, 4250 Plymouth Rd., Ann Arbor, MI 48109, USA.
Email: cladavid@med.umich.edu

the latent factors found in latent variable measurement models where multiple indicators are used to identify a latent construct. The intercept factor specifically captures status at the initial time point of data collection (although other specifications are possible), and the slope factor captures the rate of (often linear) change over time. Both latent factors vary across participants such that LGM capture both the average extent to which participants change over time and also the extent to which there are individual differences in those trajectories.

To simultaneously take advantage of the conceptual strengths of AR and LGM, several hybrid models have been developed that combine AR and LGM style parameters (Bollen & Curran, 2004). The three hybrid models we focus on here are the LCS (McArdle, 2001), ALT (Bollen & Curran, 2004), and LGM-SR (Curran et al., 2014). These models were selected as they are the most prominent hybrid autoregressive-latent growth longitudinal SEMs (Usami et al., 2019), and because they all have distinct approaches to integrating autoregressive and growth parameters (and consequently, overlapping yet distinct conceptual interpretations). Accordingly, it may not be the case that all hybrid models function similarly across different situations.

Latent change score model. The LCS captures change over time using a series of latent change score factors that reflect true but unobserved occasion-to-occasion change. A basic univariate LCS, the dual change score model (McArdle, 2001), is depicted in Figure 1. In the LCS, observed scores at each time point are separated into latent factors that capture either systematic variance in the construct of interest or error variance. The latent factors that capture systematic variance in the construct of interest are then used to specify latent change score factors that capture differences between scores at time $t-1$ and time t (lcs in Figure 1). These latent change score factors are specified as a joint function of autoregressive and growth processes so that change between time points is simultaneously determined by autoregressive and growth parameters. The autoregressive process is incorporated via a series of regression paths (β in Figure 1) that capture the extent to which scores at one time point predict more or less change between time points. The growth process is incorporated via a latent slope factor (g1 in Figure 1) that connects to latent change factors via basis coefficients (typically fixed to 1) and captures the rate of constant change over time. The residual variances of the change score factors are typically fixed to 0, making change between time points in the latent construct factors wholly a function of the two growth processes.

Autoregressive latent trajectory model. The ALT (Figure 2; Bollen & Curran, 2004, 2006) is structurally similar to the LCS (under certain conditions, the LCS is statistically subsumed by the ALT; Usami et al., 2015), but does not use latent difference score factors to model change, and observed scores are typically not separated into systematic construct and irrelevant error variance. The growth process is incorporated via latent intercept and slope factors (g0 and g1 in Figure 2) that capture a steady rate of change over time, while the autoregressive process is incorporated by adding autoregressive paths (β in Figure 2) from one observation to the subsequent observations that capture the extent to which scores at one time point predict subsequent scores. Thus, observations at each time point are specified as a function of both growth and autoregressive processes. That is, like the LCS, change over time is simultaneously determined by autoregressive and growth parameters in the ALT. Unlike the latent change score factors which have their residual variance fixed at 0, observations are not specified as completely explained by the change processes. Also, the first observation *does not* load on either the intercept or slope factor as is typical in a growth model or LCS. In Figure 2, the first observation is treated as an exogenous, predetermined variable, consistent with how the ALT model is typically presented and used in practice, though alternative specifications are possible (Jongerling & Hamaker, 2011; Ou et al., 2017).

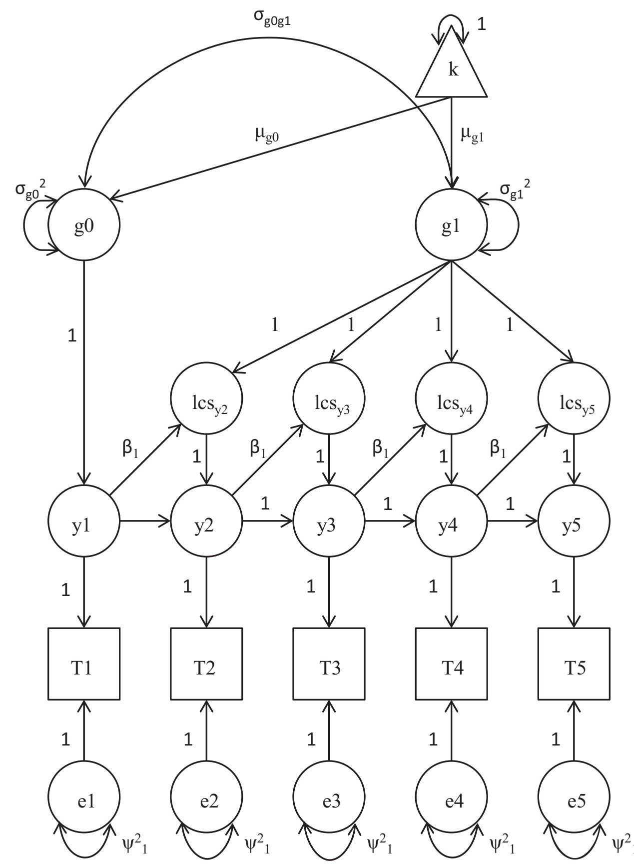

Figure 1. Latent Change Score Model. T1 . . . T5 = Time 1 through Time 5 observations; y1 . . . y5 = Time 1 through Time 5 latent variables; $lcs_{y2} \ldots lcs_{y5}$ = latent change score factors; e1 . . . e5 = Time 1 through Time 5 residuals; β = autoregressive coefficient; g0 = intercept factor; $\sigma_{g0}{}^2$ = intercept factor variance; μ_{g0} = intercept factor mean; g1 = slope factor; $\sigma_{g1}{}^2$ = slope factor variance; μ_{g1} = slope factor mean; σ_{g0g1} = intercept–slope covariance; ψ^2 = residual variance; k = constant (i.e., mean structure). Shared subscripts denote parameters commonly constrained to equality.

Latent growth model with structured residuals. In the LGM-SR (Figure 3; Curran et al., 2014), the growth process is incorporated as in a traditional LGM, with an intercept and slope factor on which all observed variables load (g0 and g1 in Figure 3). The growth model residuals are then specified as distinct latent factors in the model, and autoregressive paths (β in Figure 3) are added to this residual structure. This autoregressive process captures the extent to which deviations from the model implied trajectory at one point are related to deviations from the model implied trajectory at the next time point.

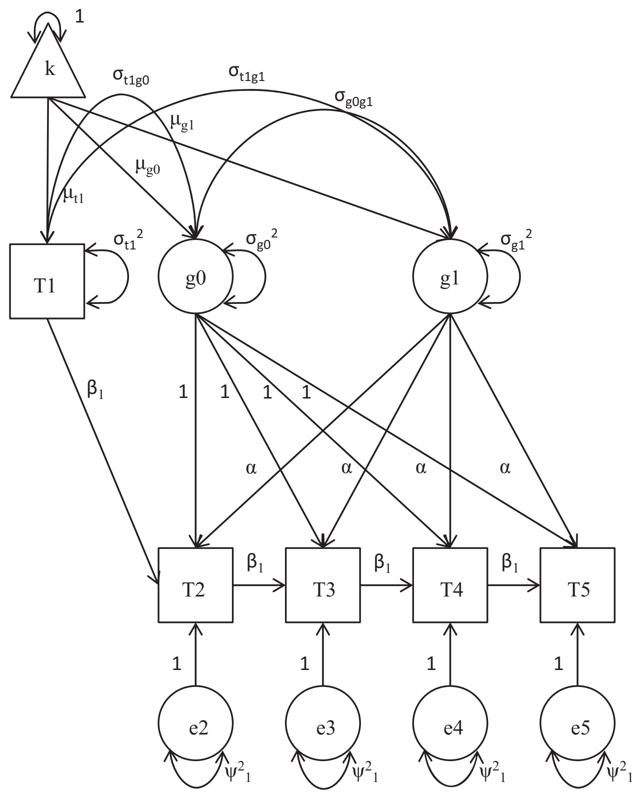

Figure 2. Autoregressive Latent Trajectory Model. T1 . . . T5 = Time 1 through Time 5 observations; e1 . . . e5 = Time 1 through Time 5 residuals; β = autoregressive coefficient; $\sigma_{t1}{}^2$ = time 1 variance; μ_{t1} = time one mean; g0 = intercept factor; $\sigma_{g0}{}^2$ = intercept factor variance; μ_{g0} = intercept factor mean; g1 = slope factor; $\sigma_{g1}{}^2$ = slope factor variance; μ_{g1} = slope factor mean; σ_{g0g1} = intercept–slope covariance; σ_{t1g0} = intercept–time one covariance; σ_{t1g1} = slope–time one covariance; α = basis coefficient; ψ^2 = residual variance; k = constant (i.e., mean structure). Shared subscripts denote parameters commonly constrained to equality.

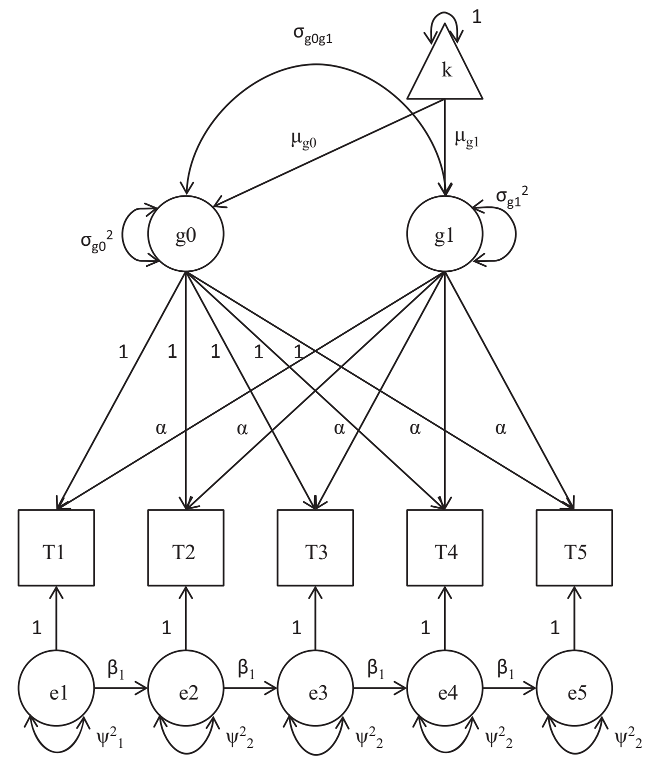

Figure 3. Latent Growth Model With Structured Residuals. T1 . . . T5 = Time 1 through Time 5 observations; e1 . . . e5 = Time 1 through Time 5 residuals; β = autoregressive coefficient; g0 = intercept factor; $\sigma_{g0}{}^2$ = intercept factor variance; μ_{g0} = intercept factor mean; g1 = slope factor; $\sigma_{g1}{}^2$ = slope factor variance; μ_{g1} = slope factor mean; σ_{g0g1} = intercept–slope covariance; α = basis coefficient; ψ^2 = residual variance; k = constant (i.e., mean structure). Shared subscripts denote parameters commonly constrained to equality.

Separating Change Processes

Hybrid models are appealing because they include both autoregressive and growth-related parameters. Empirically separating two developmental processes in a single model can be difficult in practice, however. The asymptotic correlations between autoregressive coefficients and growth factor mean estimates—two of the most critical parameters for capturing change over time—often approach $r = \pm 1$ in the LCS and ALT (Clark et al., 2018; Jongerling & Hamaker, 2011; Voelkle, 2008). Although these parameters are specified to work together to reproduce observed trajectories (Jacobucci et al., 2019), with such high correlations misspecification in one change process may be too easily accommodated by the other in service of reproducing the observed data. This can be beneficial when the goal is predicting or describing change over time (Jongerling & Hamaker, 2011; Ou et al., 2017), but as developmental researchers are typically interested in the explanation of developmental trends, these model features must be taken seriously as they could lead to erroneous conclusions.

At its core, a lack of process separability reflects a lack of unique information for effectively distinguishing between two types of change processes during parameter estimation. One of the most direct ways to provide more information along these lines is to include additional time points (Timmons & Preacher, 2015), and there is some evidence that asymptotic correlations between autoregressive coefficients and growth factor means do decrease as more time points are added to the LCS (Clark et al., 2018). This question has not yet been examined thoroughly in the LCS or other popular hybrid models though, especially in relation to robustness to misspecification. Indeed, if the inseparability of change processes during model estimation is related to the propensity to produce biased yet well-fitting solutions, and change processes become more distinguishable with more time points, hybrid models may become more robust to at least minor and routine misspecifications when there are more time points available. In longitudinal models, more time points are always preferred of course (up to a point of diminishing returns; Timmons & Preacher, 2015), but it is not always feasible to continue adding waves of assessment (Donnellan & Conger, 2007); therefore, it is critical to understand the nature and magnitude of the specific benefits gained by including more time points for distinct models.

Present Study

By integrating AR and LGM style parameters, hybrid models offer a compelling way of studying developmental dynamics. However, hybrid models often struggle to separate AR and LGM processes during estimation, which can lead to unexpected results in certain situations, such as when unrealistic parameter constraints are applied to improve convergence (Clark et al., 2018; Orth et al., 2021). The aim of the present study was to help better understand this aspect of hybrid models, specifically by examining (1) the separability of autoregressive and growth processes in three popular hybrid models (LCS, ALT, and LGM-SR), (2) if the number of time points included in these models relates to separability, and (3) if the separability between autoregressive and growth processes is related to parameter estimation and model evaluation when models are misspecified.

Method

Data were generated using Mplus version 8.1 (Muthén & Muthén, 1998–2018). All data sets contained 1,000 observations at every time point; results thus demonstrate model functioning under a "best case scenario" for longitudinal data (i.e., large sample, no attrition). For every condition described below, 1,000 unique data sets were generated and analyzed using maximum likelihood estimation. Individual Mplus input files were created and run for each simulated data set using the Mplus Automation Package (Hallquist & Wiley, 2014) in R (R Development Core Team, 2016).

Data Generation

Population models were specified in accordance with the structures depicted in Figures 1 to 3. The population parameters used in the LCS conditions come from the analysis of verbal ability in McArdle (2001), a seminal demonstration of the LCS. The population parameters in the ALT and LGM-SR conditions also came from fitting these models to the McArdle (2001) data. Across all conditions, population basis coefficients were specified in accordance with linear growth. Population model parameters and trajectories for all conditions are presented in the supplemental material (https://osf.io/ny2fw/).

Two major characteristics of the population models varied across conditions: the number of time points and the invariance of the autoregressive coefficients over time. The invariance of the autoregressive coefficients was varied when examining the functioning of misspecified models. The autoregressive coefficients were specifically varied in these conditions, as there is evidence that incorrectly constraining the autoregressive paths over time is an important form of misspecification for parameter estimation and model evaluation in hybrid models (Clark et al., 2018; Voelkle, 2008). Moreover, estimating hybrid models is often difficult in real data (Orth et al., 2021), which has resulted in the widespread application of parameter constraints—such as constraining autoregressive coefficients over time—that improve convergence but are unlikely to precisely hold in the population.

Baseline condition population models. Population models with invariant autoregressive coefficients over time are hereafter referred to as "baseline models." Data was generated from baseline models with 5, 10, 15, 20, 30, and 40 time points. In the 40 time point conditions, all parameter values were divided by 5 in order to improve computational feasibility. This did not affect the nature of the change trajectories and did not lead to results inconsistent with other conditions.

Misspecification condition population models. The population models for the misspecification conditions were equivalent to those used in the baseline conditions with the exception that autoregressive coefficients varied over time. Population models for the misspecification conditions were specified with both 5 and 20 time points. Five different misspecification conditions were considered. Four or 19 autoregressive coefficients were randomly generated from a distribution that had a mean of the original population coefficient and a standard deviation of roughly half the original population coefficient. In the first condition (MS-1), autoregressive coefficients were included in the order they were generated. In the second condition (MS-2), autoregressive coefficients were ordered from smallest to largest over time. In the third misspecification condition (MS-3), autoregressive coefficients were ordered from largest to smallest over time. In the fourth misspecification condition (MS-4), all of MS-1's original autoregressive coefficients were doubled. In the fifth misspecification condition (MS-5), all of the MS-1's original autoregressive coefficients were halved.

Data Analytic Strategy

Models that imposed invariance on the autoregressive coefficients over time were fit to the generated data. Although these constraints are not inherently part of any model considered here and can be empirically evaluated, they are common in the literature as they improve interpretability and convergence (Clark et al., 2018). The empirical asymptotic correlation matrix was obtained by taking the correlations between different parameter estimates across replications. We focus on the correlation between the autoregressive coefficient and slope mean estimates below because these parameters are the most directly relevant for representing change over time, and because past work has shown that this correlation is uniquely large in certain hybrid models but decreases with more time points (Clark et al., 2018). Full asymptotic correlation matrices can be found in the online supplemental however (https://osf.io/ny2fw/).

The average parameter estimates across replications, standard deviations across replications, average estimate of the standard error across replications, and percentage of replications with a statistically significant coefficient at the conventional $p < .05$ level were computed. Again, we focus on autoregressive coefficient and slope mean estimates below, but the full results for each condition can be found online (https://osf.io/ny2fw/). For ease of presentation, we use the average of the population autoregressive coefficient as a reference point for the estimated autoregressive coefficient below. Under the conditions considered here, the over-constrained autoregressive coefficients would ideally approximate the average population coefficient. However, though this does provide a concise summary of model functioning, it does obfuscate occasion-specific bias; the full set of population coefficients can thus be found online for making more granular comparisons between actual and estimated autoregressive effects (https://osf.io/ny2fw/).

Finally, the average values and standard deviations across replications for a set of popular fit statistics (χ^2, RMSEA, SRMR, CFI, TLI), as well as the percentage of models evidencing "adequate" and "excellent" fit, were computed (West et al., 2012). The thresholds used for denoting adequate and excellent fit conform to those typically cited in the literature (Browne & Cudeck, 1993; Hu & Bentler, 1999; West et al., 2012). Specifically, for the χ^2, the percentage of models that demonstrated significant misfit at both the .05 and .01 level was computed; for the RMSEA and SRMR, the percentage of models with values below .08 and .05 was computed; for the CFI and TLI, the percentage of models with values above .90

Table 1. Correlations Between Autoregressive Coefficients and Slope Factor Means, and Model Size, Across Baseline Conditions.

	5 Time points	10 Time points	15 Time points	20 Time points	30 Time points	40 Time points
LCS						
$r(\beta,\mu_{sl})$	−.99	−.91	−.64	−.12	<.01	−.03
# Estimated parameters	7	7	7	7	7	7
Degrees of freedom	13	58	128	223	488	853
ALT						
$r(\beta,\mu_{sl})$	−.86	−.82	−.71	−.64	−.53	−.21
# Estimated prameters	11	11	11	11	11	11
Degrees of freedom	9	54	124	219	484	849
LGM-SR						
$r(\beta,\mu_{sl})$	.04	.02	.02	.01	.05	.01
# Estimated parameters	8	8	8	8	8	8
Degrees of freedom	12	57	127	222	487	852

Note. LCS = latent change score model; ALT = autoregressive latent trajectory model; LGM-SR = latent growth model with structured residuals; $r(\beta,\mu_{sl})$ = correlation between autoregressive coefficient and slope factor mean; # Estimated parameters = number of freely estimated parameters in the model.

Table 2. Parameter Estimates and Correlations Between Autoregressive Coefficients and Slope Factor Means for LCS Across Conditions With 5 and 20 Time Points.

	Baseline		MS-1		MS-2		MS-3		MS-4		MS-5	
	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP
β: Population value	.09	.09	.10	.09	.10	.09	.12	.09	.21	.19	.05	.05
Estimate mean	.09	.09	.26	.11	.37	.18	−.12	.01	.40	.29	.17	.06
Estimate SD	.01	<.01	.01	<.01	.02	<.01	.01	<.01	.01	<.01	.02	<.01
Mean SE	.01	<.01	.01	<.01	.02	<.01	.01	<.01	.01	<.01	.02	<.01
%p < .05	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
μ_{sl}: Population value	2.06	2.06	2.06	2.06	2.06	2.06	2.06	2.06	2.06	2.06	2.06	2.06
Estimate mean	2.05	2.06	−1.67	1.19	−4.46	−2.91	8.72	8.49	−2.93	.98	−.57	1.55
Estimate SD	.37	.03	.37	<.01	.43	.04	.27	.07	.27	.03	.45	.03
Mean SE	.38	.03	.28	<.01	.43	.03	.29	.07	.30	.04	.47	.03
%p < .05	100%	100%	99%	100%	100%	100%	100%	100%	100%	100%	21%	100%
$r(\beta,\mu_{sl})$	−.99	−.12	−.99	−.17	−.99	−.59	−.97	.01	−.97	<.01	−.99	−.55

Note. MS = misspecification conditions 1 through 5; TP = time points; SD = standard deviation; SE = standard error; %p < .05 = percentage of estimates that were statistically significant at the level of p < .05; β = autoregressive coefficient; μ_{g1} = slope factor mean; $r(\beta,\mu_{sl})$ = correlation between autoregressive coefficient and slope factor mean. Autoregressive population values represent the mean autoregressive value across time.

and .95 was computed. Selected results regarding model fit are presented below, while the full results can be found in the online supplement (https://osf.io/ny2fw/).

Results

Estimation problems were rare. In some ALT conditions, a small minority of models converged with serious estimation problems, and substantially aberrant parameter estimates that affected the primary results. These outliers were excluded from the main analyses. Replications were specifically excluded in the baseline ALT condition with 5 time points (three replications removed) and 20 time points (32 replications removed), and the MS-1 (54 replications removed) and MS-5 (298 replications removed) conditions with 5 time points, and the MS-2 condition with 20 time points (118 replications removed).

Latent Change Score Model

Autoregressive and growth processes separability. The correlation between the autoregressive coefficient and slope factor mean estimates was $r = -.99$ in the baseline condition when there were only 5 time points. This correlation consistently fell as more time points were added, dropping to $r = -.12$ when there were 20 time points and $r = -.03$ when there were 40 time points (Table 1). Across misspecification conditions, these correlations also approached unity when there were 5 time points and were considerably smaller when there were 20 time points (Table 2). The autoregressive and growth processes were thus essentially indistinguishable for the LCS with 5 time points. Separability improved however as the number of time points in the model increased, particularly after at least 20 time points were included.

Parameter estimation and model evaluation. In the misspecification conditions, estimates of the autoregressive coefficient and slope factor means were consistently, substantially biased when there were 5 time points; estimates were also usually statistically significant (Table 2). These models reproduced the population trajectories well however (i.e., the parameters were biased in such a way that the observed data was still accurately reproduced) and so fit well by conventional standards (https://osf.io/ny2fw/). In the 20-time point misspecification conditions, bias in the parameter estimates

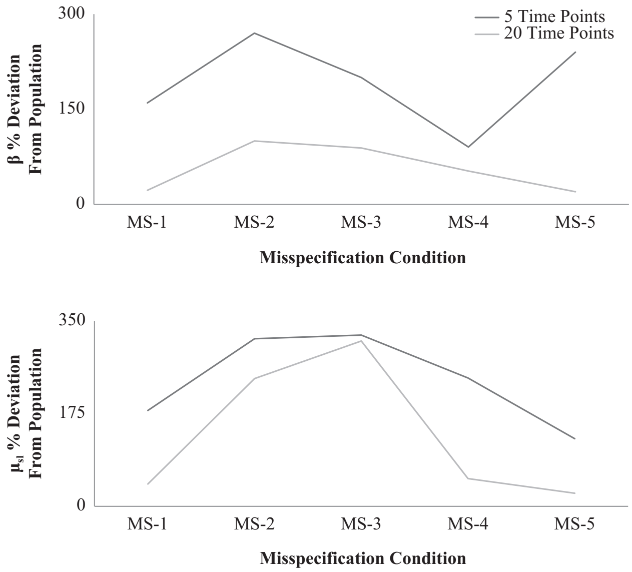

Figure 4. Percent Deviation of Average Estimated Parameter Values From Population Values Across Misspecification Conditions for the Latent Change Score Model. Results for autoregressive coefficients presented in top panel; results for slope factor means presented in bottom panel. For autoregressive coefficients, the population reference values were the average of all the individual population coefficients across time.

was less extreme; in some conditions, estimates reasonably reflected the data generating model such that the single autoregressive coefficient approximated the average population coefficient. Despite the fact that bias in the parameter estimates was reduced in the 20 time point conditions, the majority of these models were rejected on the basis of conventional fit thresholds (https://osf.io/ny2fw/).

These trends in parameter estimation and model fit are illustrated graphically in Figures 4 and 5. Figure 4 depicts how much larger or smaller the average estimated autoregressive coefficients and slope means were compared to the population values, expressed as a percentage (absolute value) of the population coefficient. This figure highlights how the estimates in the 5 time point conditions consistently deviated more from the population values than the estimates in the 20 time point conditions. However, even with 20 time points, the estimated coefficients were often at least twice (100%) as large or small as the population value. Furthermore, with both 5 and 20 time points, deviations from the population values appeared to be greater in some conditions than others, specifically MS-2 and MS-3. Figure 5 depicts the percentage of models that fit at least adequately across conditions and illustrates the asymmetry in fit such that models with 5 time points—despite their less accurate estimates—were much more likely to fit the data well compared to the models with 20 time points. The SRMR, CFI, and TLI specifically were most likely across conditions to suggest adequate fit.

Overall, when there were only 5 time points in the LCS and change processes were less separated, misspecified models demonstrated considerable bias in the parameter estimates but still fit the data well by conventional standards. Conversely, when there were 20 time points in the LCS and change processes were more distinguishable, the misspecified models demonstrated much less bias in the parameter estimates than with 5 time points, but, consistent with the misspecification, these models tended to fit the data poorly by conventional standards.

Autoregressive Latent Trajectory Model

Autoregressive and growth processes separability. The correlation between the autoregressive coefficient and slope factor mean estimates was $r = -.86$ in the correctly specified condition when there were only 5 time points (Table 1). This association gradually fell as more time points were added, dropping to $r = -.64$ when there were 20 time points, and $r = -.21$ when there were 40 time points (Table 1). Across misspecification conditions, the correlations between the autoregressive coefficient and slope factor mean estimates were large with 5 time points (rs from $-.80$ to $-.99$), and smaller with 20 time points, though still sizeable in magnitude (rs from $-.56$ to $-.79$). The autoregressive and growth processes were thus largely indistinguishable for the ALT with 5 time points. Separability did improve as the number of time points in the model increased; however, this improvement was gradual.

Parameter estimation and model evaluation. In the misspecification conditions, estimates of the autoregressive coefficients and slope factor means were considerably biased when there were 5 time points; estimates were also typically statistically significant (Table 3). These models were somewhat effective in reproducing the average population trajectories (https://osf.io/ny2fw/), and fit was often mixed across conditions. That is, in most conditions, at least some fit statistics implied an adequate or excellent model based on conventional thresholds. In the 20-time point misspecification conditions, estimates of the change parameters were usually less biased than in the 5 time point conditions, with fit again often mixed across conditions (Table 3).

These trends in parameter estimation and model fit are illustrated graphically in Figures 6 and 7. Figure 7 highlights how the estimates in the 5 time point conditions generally deviated more from the population values than the estimates in the 20 time point conditions, with some estimates more than 1000% larger or smaller than the population value. However, even with 20 time points, the estimated coefficients were often at least twice (100%) as large or small as the population value. Furthermore, with both 5 and 20 time points, deviations from the population appeared to be greater in some conditions compared to others, specifically MS-2 and MS-3 in the 20 time point conditions, and MS-2, MS-4, and MS-5 for the 5 time point conditions. Figure 7 shows how models with 5 time points were on average more likely to fit the data well compared to the models with 20 time points. The SRMR, CFI, and TLI specifically were most likely across conditions to suggest adequate fit.

Overall, when there were only 5 time points in the ALT and change processes were less separated, misspecified models demonstrated more bias compared to when there were 20 time points and processes were more separable. The fit of the models across conditions was often ambiguous, though in general the less biased models in the 20 time point conditions were somewhat more likely to be rejected on the basis of conventional thresholds for model fit.

Latent Growth Model With Structured Residuals

Autoregressive and growth processes separability. The correlations between the slope factor mean and autoregressive coefficient estimates were trivial (rs less than .10) in the baseline conditions, even with just 5 time points (Table 1). Correlations did tend to shrink as more time points were added, but this decrease was modest given the low baseline. Across misspecification conditions, the correlations between the autoregressive coefficient and slope factor mean

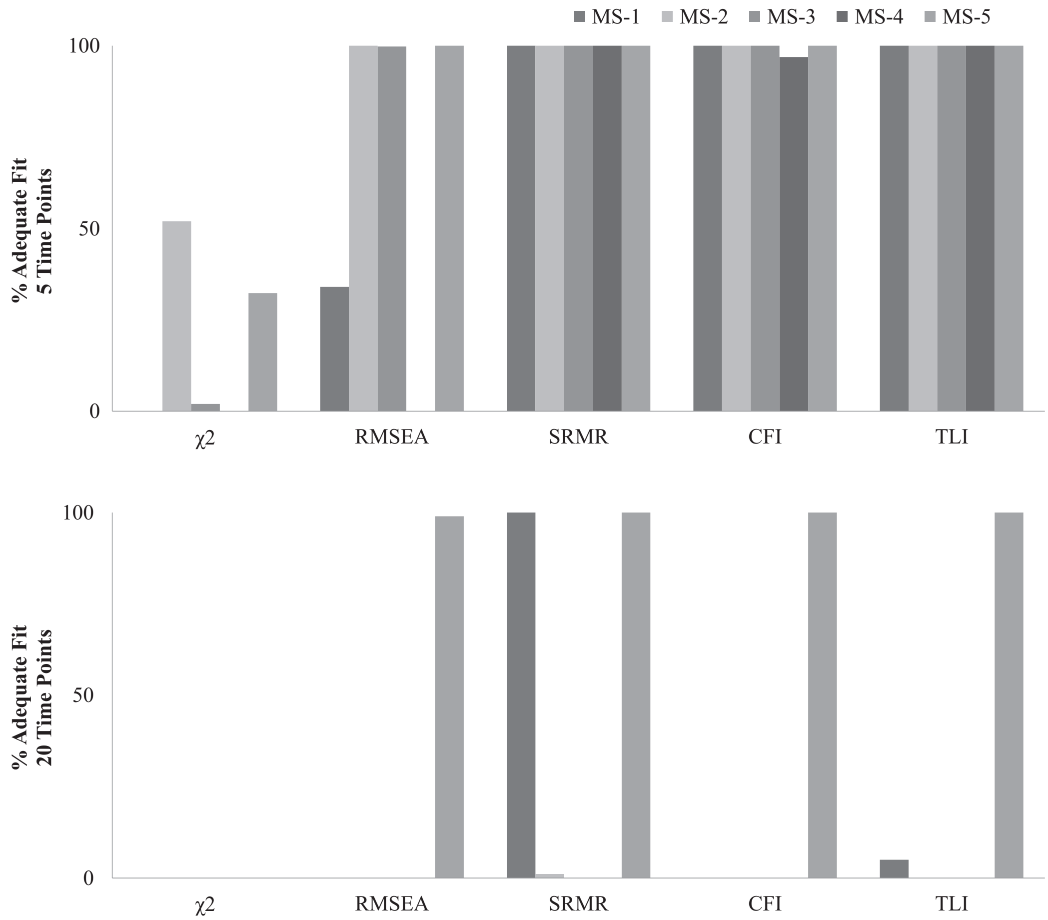

Figure 5. Percentage of Replications With Adequate Model Fit Across Misspecification Conditions for the Latent Change Score Model. Results from 5 time point conditions presented in top panel; results from 20 time point conditions presented in bottom panel. Adequate fit defined here as: χ^2 $p > .01$; RMSEA < .08; SRMR < .08; CFI > .90; TLI > .90. Complete results regarding model fit are in the online supplement (https://osf.io/ny2fw/).

Table 3. Parameter Estimates and Correlations Between Autoregressive Coefficients and Slope Factor Means for ALT Across Conditions With 5 and 20 Time Points.

	Baseline		MS-1		MS-2		MS-3		MS-4		MS-5	
	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP
β: Population value	.13	.13	.13	.13	.13	.13	.12	.13	.27	.27	.07	.07
Estimate mean	.13	.13	−.07	.15	1.46	.43	.29	.44	1.39	.30	.47	.08
Estimate SD	.03	.01	.02	.01	.02	.01	.03	.01	.02	<.01	.06	.01
Mean SE	.03	.01	.03	.01	.03	.01	.03	.01	.03	.01	.06	.01
%p < .05	99%	100%	81%	100%	100%	100%	100%	100%	100%	100%	100%	100%
μ_{sl}: Population value	5.01	5.01	5.01	5.01	5.01	5.01	5.01	5.01	5.01	5.01	5.01	5.01
Estimate mean	5.01	5.01	6.36	4.73	−2.27	4.01	2.60	2.50	−8.43	4.36	1.94	4.87
Estimate SD	.26	.07	.14	.07	.19	.06	.22	.03	.23	.06	.44	.07
Mean SE	.25	.07	.19	.07	.21	.07	.22	.05	.29	.07	.28	.07
%p < .05	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	98%	100%
$r(\beta,\mu_{sl})$	−.86	−.64	−.80	−.63	−.93	−.79	−.96	−.66	−.93	−.56	−.99	−.64

Note. MS = misspecification conditions 1 through 5; TP = time points; SD = standard deviation; SE = standard error; %p < .05 = percentage of estimates that were statistically significant at the level of $p < .05$; β = autoregressive coefficient; μ_{g1} = slope factor mean; $r(\beta,\mu_{sl})$ = correlation between autoregressive coefficient and slope factor mean. Autoregressive population values represent the mean autoregressive value across time.

estimates were similarly small in magnitude. The autoregressive and growth processes were thus largely separable for the LGM-SR regardless of the number of time points.

Parameter estimation and model evaluation. In both the 5- and 20-time point misspecification conditions, there was little bias in the parameter estimates (Table 4). To the extent there was bias, it

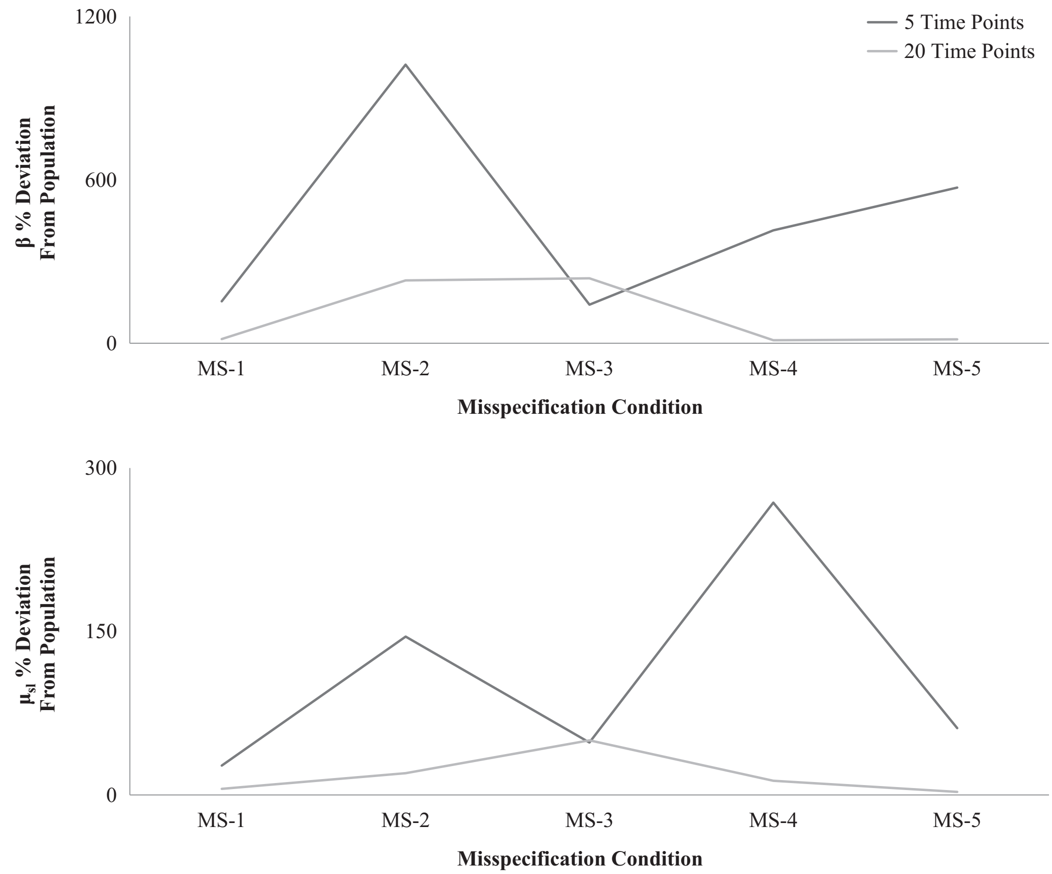

Figure 6. Percent Deviation of Average Estimated Parameter Values from Population Values Across Misspecification Conditions for the ALT. Results for autoregressive coefficients presented in top panel; results for slope factor means presented in bottom panel. For autoregressive coefficients, the population reference values were the average of all the individual population coefficients across time.

was localized to the residual structure, and minimal such that the estimate of the autoregressive coefficient was close to the average of the population coefficients. All models consequently fit the data well by conventional standards (https://osf.io/ny2fw/).

These trends in parameter estimation and model across conditions are illustrated graphically in Figures 8 and 9. Figure 8 highlights that although the estimates in the 5 time point conditions generally deviated more from the population values than the estimates in the 20 time point conditions, these deviations were usually below 50%. Furthermore, to the extent deviations were observed, they were all located in the autoregressive coefficients of the residual structure; the slope means were estimated accurately across conditions. Figure 9 illustrates how, consistent with the generally reasonable estimates, most models fit well.

Overall, with both 5 and 20 time points in the LGM-SR, change processes were separable and bias was both minimal (being about what would be expected given the nature of the misspecification) and quarantined to the parts of the model that were misspecified (i.e., a misspecified autoregressive structure did not heavily affect the slope factor). Consistent with the low degree bias, models tended to fit the data well by conventional standards.

Discussion

In this study, Monte Carlo simulation methods were used to investigate associations between change process separability, number of time points, parameter bias, and model fit, in three popular, widely used hybrid autoregressive-latent growth SEMs for longitudinal data: LCS (McArdle, 2001), ALT (Bollen & Curran, 2006), and LGM-SR (Curran et al., 2014). Overall, as more time points were included in these models, autoregressive and change process separability increased and parameter bias in the presence of misspecification decreased. This supports the notion that providing hybrid models with more raw information about change over time enables them to more effectively separate change processes during parameter estimation, and that this can potentially help improve robustness to common misspecifications.

It is difficult, however, to directly attribute the observed decreases in bias here to increased processes separability per se. For example, additional time points may simultaneously yet unrelatedly increase both separability and robustness. Furthermore, there was evidence across misspecification conditions that the underlying population trends may impact robustness to the types of misspecification considered (e.g., the LCS appeared more biased when autoregressive coefficients steadily increased or decreased over time). Still, considering the LCS, ALT, and LGM-SR individually, it appears that—even if not directly or exclusively causal—a lack of process separability at least signals a potentially considerable lack of robustness to misspecification, which warrants caution when using and interpreting these models.

Considering results across models, the LCS and ALT consistently demonstrated much less process separability than the LGM-

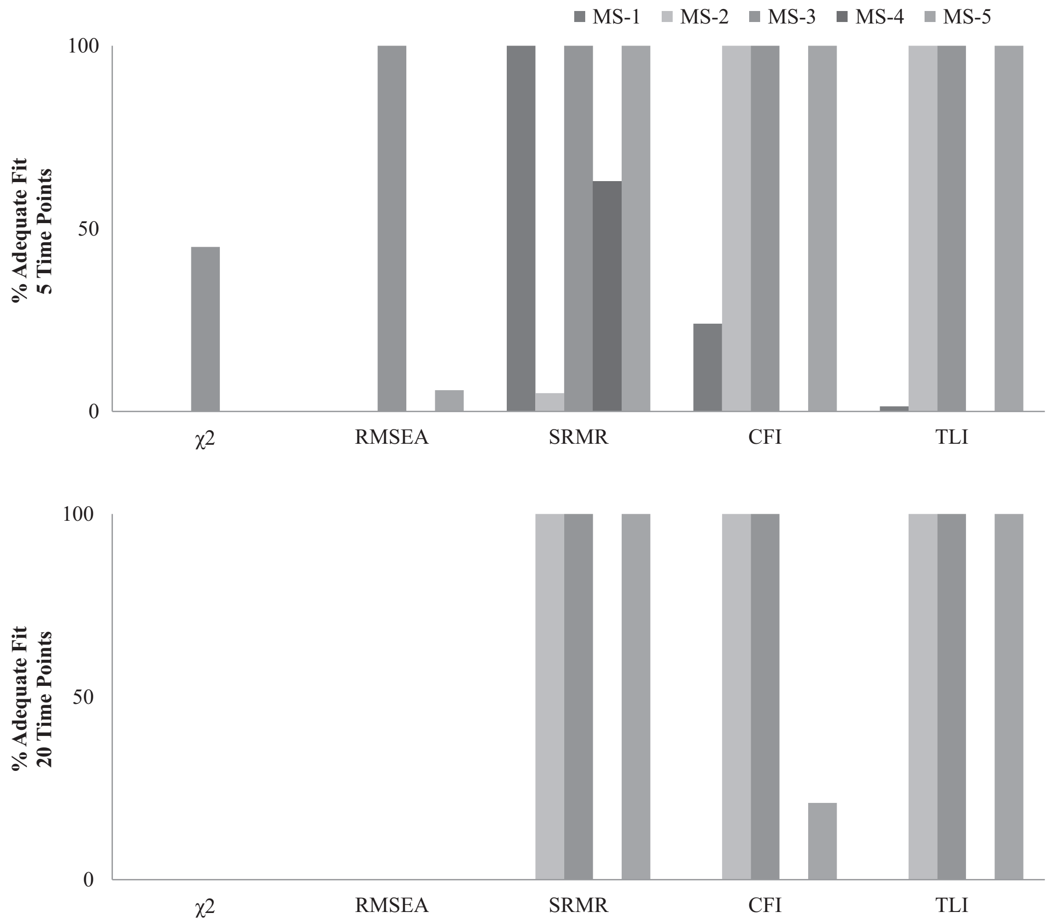

Figure 7. Percentage of Replications With Adequate Model Fit Across Misspecification Conditions for the ALT. Results from 5 time point conditions presented in top panel; results from 20 time point conditions presented in bottom panel. Adequate fit defined here as: χ^2 $p > .01$; RMSEA < .08; SRMR < .08; CFI > .90; TLI > .90. Complete results regarding model fit are in the online supplement (https://osf.io/ny2fw/).

Table 4. Parameter Estimates and Correlations Between Autoregressive Coefficients and Slope Factor Means for LGM-SR Across Conditions With 5 and 20 Time Points.

	Baseline		MS-1		MS-2		MS-3		MS-4		MS-5	
	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP	5 TP	20 TP
β: Population value	.25	.25	.25	.23	.27	.28	.26	.26	.49	.46	.12	.12
Estimate mean	.25	.25	.30	.23	.24	.28	.26	.26	.60	.45	.15	.12
Estimate SD	.04	.01	.04	.01	.04	.01	.04	.01	.05	.01	.03	.01
Mean SE	.04	.01	.04	.01	.04	.01	.04	.01	.05	.01	.03	.01
%p < .05	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99%	100%
μ_{sl}: Population value	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45
Estimate mean	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45	4.45
Estimate SD	.05	.04	.05	.04	.06	.04	.05	.04	.06	.04	.05	.04
Mean SE	.05	.04	.05	.04	.05	.04	.05	.04	.06	.04	.05	.04
%p < .05	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
$r(\beta,\mu_{sl})$	.04	.01	.05	.02	.05	.02	.04	.01	.04	.04	.05	.01

Note. MS = misspecification conditions 1 through 5; TP = time points; SD = standard deviation; SE = standard error; %p < .05 = percentage of estimates that were statistically significant at the level of $p < .05$; β = autoregressive coefficient; μ_{g1} = slope factor mean; $r(\beta,\mu_{sl})$ = correlation between autoregressive coefficient and slope factor mean. Autoregressive population values represent the mean autoregressive value across time.

SR and were more likely to produce notably biased results when the model was misspecified. These differences between the models likely follow in part from their distinct specifications, especially as the LGM-SR is not necessarily a more or less complex model (the LCS, ALT, and LGM-SR all include a similar number of free parameters, with the LGM-SR falling between the LCS and ALT;

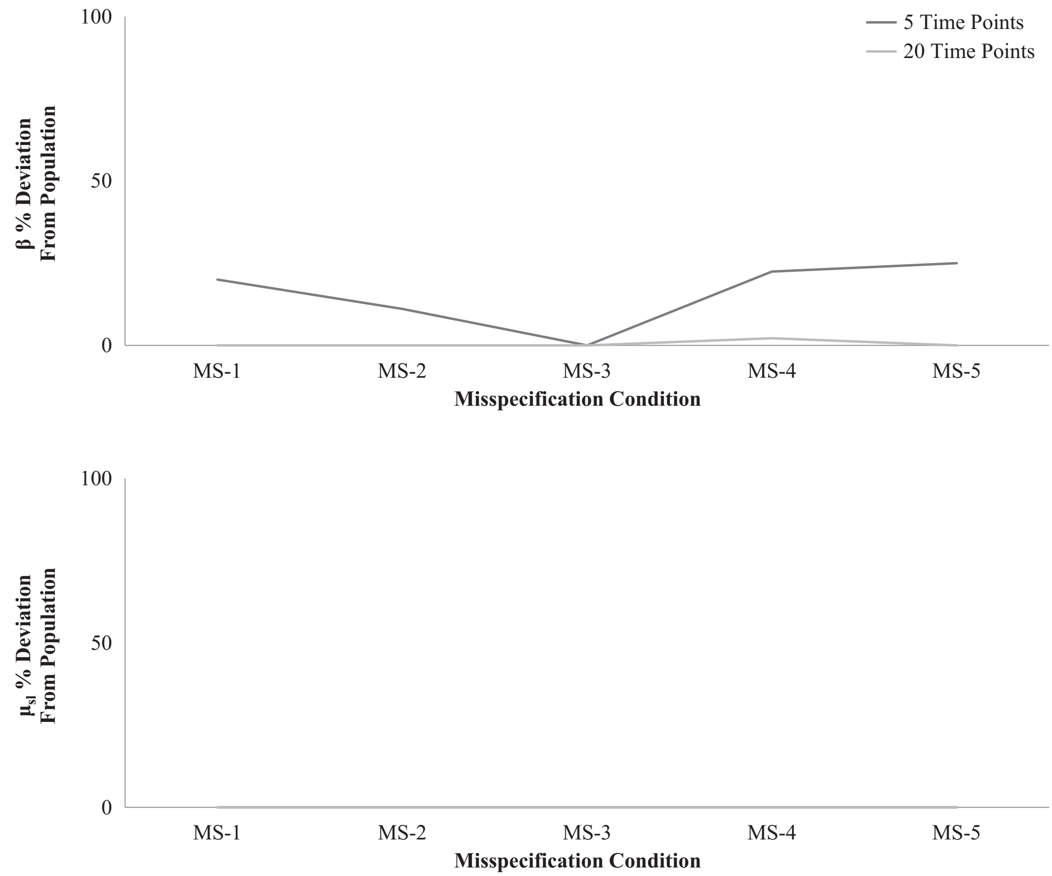

Figure 8. Percent Deviation of Average Estimated Parameter Values From Population Values Across Misspecification Conditions for the LGM-SR. Results for autoregressive coefficients presented in top panel; results for slope factor means presented in bottom panel. For autoregressive coefficients, the population reference value was the average of all the individual population coefficients across time.

Table 1). Indeed, this pattern of results maps on to the distinction Usami and colleagues (2019) recently drew between joint (or "accumulator") and non-joint hybrid models. Although all these models incorporate growth and autoregressive processes, models such as the LCS and ALT that specify outcomes as a joint function of change processes can be thought to represent a qualitatively distinct subset of models with unique interpretations and considerations. This is consistent with the present findings such that it was the two "joint" models specifically that struggled to separate change processes and that produced considerably biased results when misspecified (especially with fewer time points).

Conversely, when autoregressive and growth processes were isolated, as in the LGM-SR, asymptotic covariances between parameter estimates were small, and bias was minimal, largely regardless of how many occasions of assessment there were. This emphasizes that the number of time points per se is not sufficient to broadly explain associations between parameter estimation, model evaluation, and misspecification in hybrid models. That is, more occasions of assessment are preferable to less, but the specific advantages gained depend on other factors such as model form, and the processes operating in the population. All of these findings follow from how the models were designed, but do still carry implications for developmental researchers, suggesting that the use of certain hybrid SEM (LCS and ALT) entails inferential risks even with a larger than typical number of time points available.

Implications and Recommendations

The LCS and ALT have several major limitations. As illustrated here, these models are generally ill-equipped to accurately characterize developmental processes when only a small number of time points are available. Including more time points can address this, but including more time points is not always feasible, especially to the extent necessary to eliminate concerns (Timmons & Preacher, 2015). Models could also be estimated with as few parameter constraints as possible. However, less restricted models often fail to converge, and although even a single extra constraint can improve convergence, this usually still leads considerable bias (Clark et al., 2018). Indeed, a highly plausible scenario is one in which there are not enough time points to constrain the model without potentially introducing severe bias, yet a fully unconstrained model cannot be effectively estimated. Overall, the results here and from other methodological studies (e.g., Clark et al., 2018; Voelkle, 2008) indicate that LCS and ALT models have serious, difficult to accommodate idiosyncrasies that could readily lead to erroneous developmental conclusions. Accordingly, we *(1) strongly recommend that developmental researchers avoid applying the ALT model to their data, and exercise extreme caution with the LCS*, which although slightly more robust and responsive here exhibited similar properties as the ALT across conditions.

On the other hand, the LGM-SR performed well across conditions. It was fairly robust in the face of misspecification no matter

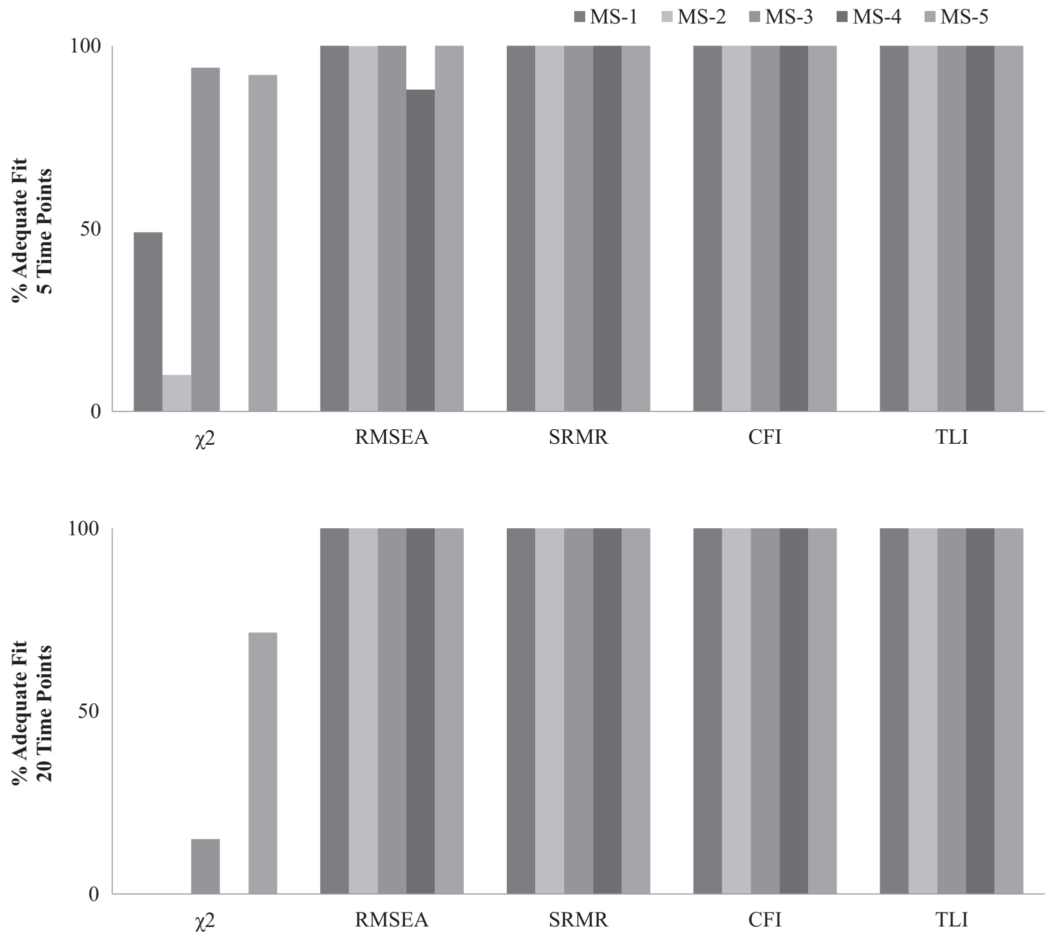

Figure 9. Percentage of Replications With Adequate Model Fit Across Misspecification Conditions for the LGM-SR. Results from 5 time point conditions presented in top panel; results from 20 time point conditions presented in bottom panel. Adequate fit defined here as: χ^2 $p > .01$; RMSEA < .08; SRMR < .08; CFI > .90; TLI > .90. Complete results regarding model fit are in the online supplement (https://osf.io/ny2fw/).

how many time points were included, and parameter estimates were only weakly correlated. To be sure, by isolating the two change processes, the LGM-SR may appear less conceptually compelling than the LCS and ALT. The LGM-SR is indeed a less literal synthesis of autoregressive and growth processes, and its time-specific residual factors exclusively capture within-person variance, which may not always be ideal when time varying covariates are considered. However, the LGM-SR's isolation of change processes clearly has some practical advantages. Model selection should of course ultimately be guided by theoretical concerns, but in the context of hybrid models, the LCS and ALT are too limited to confidently apply, even if they do represent a close match to a theoretical question.

The results for the LGM-SR show that growth and autoregressive processes can be simultaneously incorporated into a single model without becoming seriously entangled. In other words, the limitations of the LCS and ALT are not inherent limitations of hybrid models per se. Rather, it appears that it is the specific method of combining autoregressive and growth processes that has the most implications for model performance. Thus, we *(2) also recommend that researchers interested in simultaneously modeling growth and autoregressive processes in an SEM framework begin with the LGM-SR or similar models that isolate the two distinct change processes*. For example, one promising alternative to the LGM-SR is the Random Intercept–Cross Lagged Panel Model (RI-CLPM; Hamaker et al., 2015). The RI-CLPM is equivalent to a LGM-SR with 0 slope factor variance and includes a latent intercept factor that captures stable, trait-like between person variation across the study. The RI-CLPM has many benefits, including that it can be identified with fewer waves of data than the LGM-SR, the potentially appealing interpretation of the random intercept, and how it provides a parsimonious method of modeling longitudinal trends when there is little change and/or individual variability in change over the course of the study (e.g., when studies are short and/or focus on rare behaviors). Although various properties of the LGM-SR and RI-CLPM still need to be investigated, the current results imply that models specified along these lines represent a more reliable approach for considering different types of trends over time.

Notably, only univariate models were considered here, but univariate models are rarely the primary models of interest. That is, researchers are most often interested in the inclusion of time invariant and time varying covariates in order to predict developmental trends. There is evidence that paths from time invariant predictors to the slope factors in LCS will also be severely biased under conditions like those considered here (Clark et al., 2018). However, as noted above, the limitations observed in the univariate LCS and ALT suggest that they should be avoided in general, and not extended with multiple variables. The results for the LGM-SR were more encouraging, suggesting that the LGM-SR and related models like the RI-CLPM are more viable approaches for examining predictors of change and bidirectional influences over time. More work is needed to confirm this, however. Also, it is important to highlight here that a major goal in the development of the LGM-SR and RI-CLPM was the disaggregation of within and between person effects. The residual structure thus specifically captures within-person trends over time, which is distinct from standard time varying covariates growth models, and how time varying covariates are typically included in LCS and ALT models.

Future Directions

Examining the LGM-SR and similar models across bi- and multivariate contexts is a critical next step. This raises the issue of separating growth, autoregressive, *and* cross-lagged processes in estimation, the latter two of which are not as separated in the model specification as the growth and autoregressive processes. Closer investigations of the LGM-SR, RI-CLPM, and their interrelations are also warranted. For example, to what extent can slope factor variability be ignored in an RI-CLPM before parameter estimates in the residual structure begin to diverge meaningfully from the LGM-SR (especially in shorter studies where the RI-CLPM is more feasible)? Finally, the current study included large samples, no attrition, and evenly spaced assessment intervals. These are idealized data conditions that most actual studies will deviate from. Future work can thus more thoroughly examine model functioning under more realistic data conditions. However, real-world data conditions will only exacerbate the challenges we identified in models applied to ideal data.

Conclusion

Hybrid autoregressive-growth longitudinal models offer both conceptual and practical benefits. This study highlighted, however, that in many common situations some of these models struggle to distinguish between different change processes during estimation. Although the LGM-SR consistently separated the different change processes during estimation, the LCS and ALT both struggled until many time points were included in the model, which are likely not practical in real-world longitudinal studies. The LCS and ALT were also more likely to provide extremely biased but well-fitting solutions, potentially leading to inaccurate conclusions. Broadly speaking the LGM-SR appears to be the safest, most robust hybrid autoregressive-growth longitudinal model across contexts.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Address correspondence to D. Angus Clark, Department of Psychiatry, Addiction Center, University of Michigan, Rachel Upjohn Building, 4250 Plymouth Rd., Ann Arbor, MI 48109. Electronic mail may be sent to cladavid@med.umich.edu. A portion of this project was presented as a poster at the 2018 Developmental Methods Conference in Whitefish, Montana.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305A110293 and R324A150063 awarded to Michigan State University, and the National Institute on Alcohol Abuse and Alcoholism through Grant T32 AA007477 (F. Blow).

ORCID iDs

D. Angus Angus Clark https://orcid.org/0000-0002-8334-492X
Amy K. Nuttall https://orcid.org/0000-0003-3382-3516

Supplemental Material

Supplemental material for this article is available online.

References

Biesanz, J. C. (2012). Autoregressive longitudinal models. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 459–471). Guilford Press.

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*. John Wiley & Sons.

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Clark, D. A., Nuttall, A. K., & Bowles, R. P. (2018). Misspecification in latent change score models: Consequences for parameter estimation, model evaluation, and predicting change. *Multivariate Behavioral Research*, *53*(2), 172–189.

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.

Donnellan, M. B., & Conger, R. D. (2007). Designing and implementing longitudinal studies. R. W. Robins, R. C. Fraley, & R. Krueger (Eds.). *The Handbook of research methods in personality psychology*. (pp. 21–36). Guilford Press.

Hallquist, M., & Wiley, J. (2014). MplusAutomation: Automating Mplus model estimation and interpretation. R package version 0.6-3.

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

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Jacobucci, R., Serang, S., & Grimm, K. J. (2019). A short note on interpretation in the dual change score model. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 924–930.

Jongerling, J., & Hamaker, E. L. (2011). On the trajectories of the predetermined ALT model: What are we really modeling? *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(3), 370–382.

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. In S. DuToit, R. Cudeck, & D. Sorboom (Eds.), *Structural equation modeling: Present and future* (pp. 341–380). Scientific Software International, Inc.

Muthén, L. K., & Muthén, B. O. (1998–2018). *Mplus user's guide*. 8th Ed. Muthén & Muthén.

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. http://dx.doi.org/10.1037/pspp0000358

Ou, L., Chow, S. M., Ji, L., & Molenaar, P. C. (2017). (Re) evaluating the Implications of the autoregressive latent trajectory model through likelihood ratio tests of its initial conditions. *Multivariate Behavioral Research*, *52*(2), 178–199.

R Core Team. (2016). R: A language and environment for statistical computing. (3.3.1) [computer program]. Vienna, Austria: R Foundation for Statistical Computing.

Timmons, A. C., & Preacher, K. J. (2015). The importance of temporal design: How do measurement intervals affect the accuracy and efficiency of parameter estimates in longitudinal research? *Multivariate Behavioral Research*, *50*(1), 41–55.

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., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657.

Voelkle, M. C. (2008). Reconsidering the use of autoregressive latent trajectory (ALT) models. *Multivariate Behavioral Research*, *43*(4), 564–591.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In *Handbook of structural equation modeling* (pp. 209–231). The Guilford Press.