

Including Time-Invariant Covariates in the Latent Growth Curve Model

Reinoud D. Stoel and Godfried van den Wittenboer
University of Amsterdam

Joop Hox
Utrecht University

Within the latent growth curve model, time-invariant covariates are generally modeled on the subject level, thereby estimating the effect of the covariate on the latent growth parameters. Incorporating the time-invariant covariate in this manner may have some advantages regarding the interpretation of the effect but may also be incorrect in certain instances. In this article we discuss a more general approach for modeling time-invariant covariates in latent growth curve models in which the covariate is directly regressed on the observed indicators. The approach can be used on its own to get estimates of the growth curves corrected for the influence of a 3rd variable, or it can be used to test the appropriateness of the standard way of modeling the time-invariant covariates. It thus provides a test of the assumption of full mediation, which states that the relation between the covariate and the observed indicators is fully mediated by the latent growth parameters.

Since its introduction in structural equation modeling (SEM) by McArdle (1986, 1988) and Meredith and Tisak (1984, 1990), latent growth curve (LGC) modeling has seen an increasing number of applications in social and behavioral sciences. Also, much literature has been devoted to the technical development of the LGC model and its application to substantive research questions (e.g., Duncan, Duncan, Strycker, Li, & Alpert, 1999; Hox, 2002; McArdle & Bell, 2000; Muthén & Khoo, 1998; Willett & Sayer, 1994). One of the purposes of LGC modeling is to relate growth parameters to individual characteristics, background, and environmental factors, to detect systematic interindividual differences in the individual growth

Requests for reprints should be sent to Reinoud D. Stoel, Department of Biological Psychology, Vrije Universiteit Amsterdam, Van der Boechorstraat 1, 1081 BT Amsterdam, The Netherlands. E-mail: rd.stoel@psy.vu.nl

curves. These time-invariant covariates, or individual-level predictors, are generally modeled on the subject level. A direct effect of the covariate is specified on the parameters describing the growth curve (i.e., [initial] level and shape), which enables one to investigate whether the covariate explains part of the interindividual differences in the growth curves.

A disadvantage of modeling time-invariant covariates in this manner is that it is based on the assumption that the relation between the covariate and the observed indicators is fully mediated by the growth parameters. We refer to this assumption as the *assumption of full mediation*. That is, it is assumed that the direct effect of the time-invariant covariate on each of the indicators is equal to zero. The model with this assumption relaxed (i.e., both types of effects specified) is not identified. Consequently, if there are direct effects of the time-invariant covariate on the indicators in the population, this standard way of modeling time-invariant covariates will not be correct. Estimating the standard model in such situations leads to biased estimates of the growth parameters, and erroneously high values of the chi-square test of model fit. In this article we discuss an alternative approach for modeling time-invariant covariates in LGC analysis that has not received much attention yet in the literature. Moreover, we will show that the standard way of modeling time-invariant covariates constitutes a special case of this approach.

A LATENT GROWTH CURVE MODEL WITH TIME-INVARIANT COVARIATES

In general, time-invariant covariates are modeled on the subject, or factor level, of the LGC model (e.g., Willett & Sayer, 1994), as presented in the Equations 1 to 3 of Model 1

$$y_{it} = \alpha_t + \lambda_0 \eta_{0i} + \lambda_{1t} \eta_{1i} + \varepsilon_{it} \quad (1)$$

$$\eta_{0i} = \nu_0 + \beta_0 z_i + \zeta_{0i} \quad (2)$$

$$\eta_{1i} = \nu_1 + \beta_1 z_i + \zeta_{1i} \quad (3)$$

in which z_i represents the time-invariant covariate. It is used here as a predictor of the latent level factor, η_{0i} , and latent shape factor, η_{1i} , with regression parameters of, respectively, β_0 and β_1 . The parameters ν_0 and ν_1 represent the means of, respectively, the latent level and latent shape factor; λ_0 is the basis function for the level factor, and all its values are set equal to 1; λ_{1t} , the basis function for the shape factor, represents scores related to time (e.g., measurement occasion or age) and may be fixed or partially estimated to represent nonlinear growth. Residual deviations from the latent level and shape factor are represented by, respectively, ζ_{0i} and

ζ_{1t} with variances ψ_{00} and ψ_{11} , and ε_{it} is a time-specific residual with variance $\sigma^2_{\varepsilon t}$. The intercepts of the indicators are represented by α_t . Because this model consists of just one indicator per time point, α_t will be fixed to zero. A linear variant of Model 1 is represented as a path diagram in Figure 1, for four measurement occasions, implying that λ_{1t} equals [0, 1, 2, 3]. We refer to the references mentioned in the first paragraph of this article for a detailed explanation of the LGC model.

Model 1, which we entitle the *growth predictor* model, presents the standard approach to the modeling of time-invariant covariates in LGC modeling, as well as in multilevel regression (MLR) analysis (Bryk & Raudenbush, 1987, 1992; Goldstein, 1987, 1995). An advantage of Model 1 is that the effect of the covariate on the growth parameters can be inferred directly, and the proportion of variance of the latent growth parameters explained by the covariate is produced by the software, or it can be computed easily by hand. This is an attractive property of the growth predictor model because it corresponds to the standard practice of current statistical modeling as well as to the theoretical interest of many researchers. Often, our theories focus on factors to explain growth and the interindividual variability in growth, and not in a description of latent growth curves per se. Modeling the effect of time-invariant covariate on the latent level is a logical extension of the LGC model from such perspective.

However, as noted previously, modeling a time-invariant covariate with direct effects on the parameters describing the growth curve amounts to a very restricted model. This model makes the assumption of full mediation. That is, it is assumed that the direct effects of the covariate on the residual variances of the indicators is

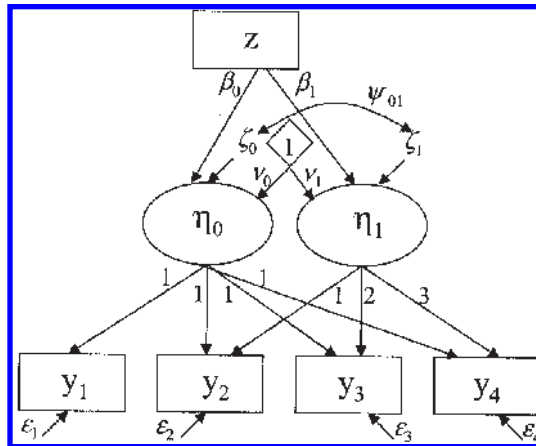


FIGURE 1 Growth predictor growth curve model (time-invariant covariate modeled on the latent level). Latent factor intercepts are conceptualized as regression on a constant equal to one (see Hancock et al., 2001). The curved double-headed arrow represents the correlation between the latent factors.

equal to zero. However, although the covariate is invariant over time, the effect of the covariate on each of the indicators does not necessarily need to follow the pattern forced by the basis function of the growth predictor model. In other words, if the latent growth parameters do not fully mediate the effect of the time-invariant covariate on the indicators, the covariate can be related to some of the residual variances of the indicators. If this is the case, Model 1 constitutes a misspecified model because the direct effects of the covariate on the indicator variables are not specified. In an extreme case it might even be that the latent growth parameters do not mediate the effect of the covariate at all. In that case, the covariate may have a direct effect only on the indicators at some, but not all, of the time points. Such differential effects cannot be incorporated in Model 1 because the effect of z_i is modeled on the latent level. Estimating Model 1 when the effect of the covariate does not follow the pattern forced by the basis function of the growth predictor model will lead to biased estimates of some of the model parameters and goodness-of-fit measures, or it will cause convergence problems.

Fortunately, the growth predictor model is not the only way a time-invariant covariate can be integrated in an LGC model. Time-invariant covariates can also be modeled with direct effects on the indicators at each occasion as presented by Equations 4 to 6 of Model 2, and Figure 2:

$$y_{it} = \alpha_t + \lambda_0 \eta_{0i} + \lambda_{1t} \eta_{1i} + \gamma_t z_i + \varepsilon_{it} \tag{4}$$

$$\eta_{0i} = v_0 + \zeta_{0i} \tag{5}$$

$$\eta_{1i} = v_1 + \zeta_{1i} \tag{6}$$

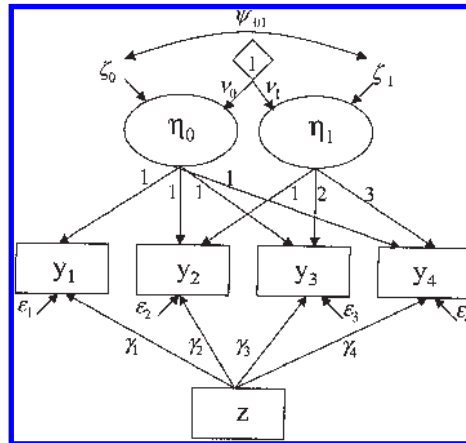


FIGURE 2 Direct effect growth curve model (time-invariant covariate modeled directly on the observed indicators).

where γ_t (having a time-specific subscript t) represents the effects of z_i on each of the t indicators. Assuming that at least one of the γ_t 's is significant, estimating the model of Equations 4 to 6 implies that the growth parameters are estimated, after the effect of the time-invariant covariate has been controlled for, or partialled out. To put it differently, in Model 2, which shall be entitled as the *direct effect* model, the growth curve will be estimated after being corrected for the effects of the time-invariant covariate.¹ In addition, where in the growth predictor model the covariate could only account for the variances of the latent level and shape factor, respectively, ψ_{00} and ψ_{11} , it can now also account for the time-specific variance $\sigma^2_{\epsilon t}$ at each occasion. This may be important because if a relation of this type is present between a time-invariant covariate and the residual variances, omitting these relations from the model might lead to an unfair rejection of the model.

The Relation Between the Growth Predictor and the Direct Effect Model

Although it may not be obvious at first sight, the growth predictor model (Model 1) constitutes a special case of the direct effect model depicted in Model 2, as is shown in Appendix A. Extra restrictions put on the direct effect model lead to the growth predictor model, so that the growth predictor model is nested within the direct effect model. The restrictions imposed on the direct effect model by assuming that the covariate has an effect on the latent factor can therefore be tested explicitly using a likelihood ratio test.

Although the direct effect model will give unbiased estimates of the growth parameters, computation of both the effect of the time-invariant covariate on the growth parameters and the variance explained is more problematic because these parameters may be confounded with the direct effects on the indicators. Here, a distinction must be made between two cases: (a) The likelihood ratio test to test the more restricted growth predictor model presented in Equations 1 to 3 against the direct effect model presented in Equations 4 to 6 is not significant (i.e., the assumption of full mediation holds), and (b) the likelihood ratio test is significant (i.e., violation of the assumption of full mediation).

The first case poses no special problems. The restrictions imposed by the model can be accepted and the effects of the covariate on the latent growth parameters can be interpreted, as well as their explained variance. However, in the second case, where the likelihood ratio test indicates that the more restricted growth predictor model has to be rejected, the assumption of full mediation is violated, and the model needs to be estimated using the direct effect model. Any effect of the time-invariant covariate on the latent growth parameters will now be confounded

¹Compare this to controlling for the effect of a covariate as defined in analysis of covariance (ANCOVA; cf. Rovine & Molenaar, 2001).

with the direct effects, γ_t^* , of the covariate on the indicators. Given the basis function, λ_{1t} , for the growth rate factor, the estimated effect, γ_t , of the time-invariant covariate on the observed indicator at occasion t , will be

$$\gamma_t = \lambda_0 \beta_0 + \lambda_{1t} \beta_1 + \gamma_t^* \quad (7)$$

Equation 7 shows that the effect of the covariate on the indicator at time t now consists of three parts. One part is due to the effect of the covariate on the latent level factor, the second part is due to the effect of the time-invariant covariate on the latent growth rate factor, and the third part is the direct effect of the covariate on occasion t .

Substituting Equation 7 into the direct effect model gives Model 3,

$$y_{it} = \alpha_t + \lambda_0 \eta_{0i} + \lambda_{1t} \eta_{1i} + \gamma_t^* z_i + \varepsilon_{it} \quad (8)$$

$$\eta_{0i} = \nu_0 + \beta_0 z_i + \zeta_{0i} \quad (9)$$

$$\eta_{1i} = \nu_1 + \beta_1 z_i + \zeta_{1i} \quad (10)$$

which can be regarded as a combination of Models 1 and 2. Model 3 is not identified without additional constraints and cannot be estimated as such. Thus, the time-invariant covariate has to be treated as a true covariate, in the sense that no inferences are made about its predictive value for the latent growth parameters.

Attractive solutions might be to either estimate the growth predictor model, and thus constrain all γ_t^* parameters to zero, or estimate the model using the direct effect model, and thus constrain β_0 and β_1 to zero, and to compute the explained variance by hand (e.g., Willett & Sayer, 1994). However, as illustrated in the next section, both solutions may lead to biased estimates of the respective explained variances because both solutions use information from misspecified models.

EXAMPLE

In this section an illustration is given of the approach to model the effect of time-invariant covariates directly on the observed variables (Model 2: the direct effect model). The approach is compared to the traditional way of modeling the effect of the time-invariant covariate on the latent variables (Model 1: the growth predictor model). Two artificial data sets are generated with Prelis 2.3 consisting of 300 subjects (i), measured on 4 occasions (y_{it}), and one time-invariant covariate (z_i). The data are generated such that they fit perfectly to the true model, being (a) an LGC model with effects of covariate on the latent level and shape factor, or (b) an LGC model with effects of the covariate on the observed variables of the first two occasions. Subsequently, these data sets are analyzed twice using models corresponding to Models 1 and 2. Thus, a total of four models are analyzed and the results are compared pairwise. The data analyses are performed using the maximum likelihood estimation procedure of LISREL 8.51. We refer to Appendixes B and C at the end of this article for the corresponding covariance matrices and mean vectors. Ta-

bles 1 and 2 present chi-square goodness-of-fit measures and parameter estimates. The second column of Tables 1 and 2 presents the population values of the true model according to which the data have been generated.

Table 1 presents the case in which the data have been generated with the effect of the covariate on the latent level. It can be seen that, as expected, the estimates of the parameters of the true model (the growth predictor model, Model 1a) match perfectly with the population parameters. This model fits perfectly to the data, $\chi^2(7) = .00, p = 1.00$. However, the model with the covariate modeled directly on the observed variables (the direct effect model, Model 2a) also provides a perfect fit to the data, be it with 2 *df* less, $\chi^2(5) = .00, p = 1.00$. As a result, estimates of the growth parameters and test statistics are the same. Furthermore, it can be seen that the effect of the covariate on the observed variables increases over time from $\gamma_1 = 1$ to $\gamma_4 = 4$. This is in correspondence with the pattern forced by the basis function for the level factor, [1, 1, 1, 1], and the shape factor [0, 1, 2, 3]. The likelihood ratio test, $\chi^2(7 - 5) = .00, p = 1.00$, does not lead to a rejection of the more restricted Model 1a, which is in fact the true model.

The important message from Table 1 is that both models can recover the population parameters when the covariate has an effect that is in correspondence with the pattern forced by the basis functions for the level and shape factors. For reasons of parsimony and ease of interpretation, modeling the covariate on the latent variable might be preferred in this situation, as discussed before.

However, in some situations the two models do not lead to the same parameter estimates and model fit. Table 2 presents the case in which the covariate has an ef-

TABLE 1
Analysis of the Data Generated With an Effect of the Covariate on the Latent Variable. Model 1a Represents the True Model

<i>Parameter</i>	<i>Population Value</i>	<i>Model 1a</i>	<i>Model 2a</i>
Fixed part			
v_0	10	10.00 (174.4)	10.00 (174.4)
v_1	2	2.00 (41.7)	2.00 (41.7)
β_0	1	1.00 (17.4)	
β_1	1	1.00 (20.8)	
γ_1	0		1.00 (16.8)
γ_2	0		2.00 (26.5)
γ_3	0		3.00 (27.3)
γ_4	0		4.00 (26.52)
Random part			
$\sigma^2_{\epsilon t}$.25	.25/ .25/ .24/ .24	.25/ .25/ .24/ .24
Ψ_{00}	.81	.81 (9.7)	.81 (9.7)
Ψ_{11}	.64	.64 (11.3)	.64 (11.3)
Ψ_{01}	0	.00 (.0)	.00 (.0)
χ^2		$\chi^2(7) = .00; p = 1.00$	$\chi^2(5) = .00; p = 1.00$

Note. Test statistics (est./SE) are given in parentheses.

TABLE 2
 Analysis of the Data Generated With an Effect of the Covariate
 on the Observed Variables at the First Two Occasions. Model 2b
 Represents the True Model

<i>Parameter</i>	<i>Population Value</i>	<i>Model 1b</i>	<i>Model 2b</i>
Fixed part			
v_0	10	10.00 (155.5)	10.00 (174.3)
v_1	2	2.00 (40.7)	2.00 (41.7)
β_0	0	.54 (9.4)	
β_1	0	-.19 (-4.0)	
γ_1	.5		.5 (8.4)
γ_2	.5		.5 (6.6)
γ_3	0		.00 (.0)
γ_4	0		.00 (.0)
Random part			
$\sigma^2_{\epsilon t}$.25	.24/ .29/ .31/ .18	.25/ .25/ .25/ .25
ψ_{00}	.81	.81 (9.6)	.81 (9.6)
ψ_{11}	.64	.65 (11.3)	.64 (11.2)
ψ_{01}	0	-.00 (-.1)	-.00 (-.0)
χ^2		$\chi^2(7) = 55.35; p = .00$	$\chi^2(5) = .00; p = 1.00$

Note. Test statistics (est./SE) are given in parentheses.

fect on just the first two occasions. Although this might be considered an extreme case that does not occur in reality, it represents a clear violation of the assumption of full mediation, and it illustrates the possibilities of the alternative direct effect model proposed here.

Model 2b, which is the true model, perfectly recovers the population parameters of both the growth parameters and the effects of the covariate. Model 1b, however, does provide correct estimates of the means, variances, and covariance of the latent level and shape factors, but the estimated effects of the covariate are not correct. According to Model 1b, the time-invariant covariate explains, respectively, 27% and 5% of the intercept and slope factors, whereas these effects are in fact equal to zero in the population. At the same time the effects of the covariate on the observed variables given the basis function for the level and shape factor of Model 1b are computed by hand as, respectively, $\gamma_1 = .54$, $\gamma_2 = .35$, $\gamma_3 = .16$, $\gamma_4 = -.03$. These effects are, as expected, also unequal to the population parameters according to which the data have been generated. In addition, the fit of the model is bad, $\chi^2(7) = 55.35$, $p = .00$. As a consequence, the likelihood ratio test, $\chi^2(7 - 5) = 55.35$, $p = .00$, leads to a rejection of Model 1b in favor of Model 2b. A comparison of the variances of the growth parameters of the direct effect model (Model 2b) with the estimates of an LGC model without the time-invariant covariate as presented in Table 3 (Model 4) to compute the explained variance gives similar results as previously shown. Although in the population the effect of the covariate on the

TABLE 3
 Analysis of the Same Data as Table 2. Model 4 Represents an LGC Model
 Without the Time-Invariant Covariate

<i>Parameter</i>	<i>Population Value</i>	<i>Model 4</i>
Fixed part		
v_0	10	9.99 (153.4)
v_1	2	2.00 (40.7)
β_0	0	
β_1	0	
γ_1	.5	
γ_2	.5	
γ_3	0	
γ_4	0	
Random part		
$\sigma^2_{\epsilon t}$.25	.17/ .33/ .31/ .16
Ψ_{00}	.81	1.13 (10.3)
Ψ_{11}	.64	.69 (11.5)
Ψ_{01}	0	-.12 (-2.1)
χ^2		$\chi^2(5) = 8.46; p = .13$

Note. Test statistics (est./SE) are given in parentheses.

growth parameters is equal to zero, this comparison gives, respectively, 28% and 7% explained variance for the intercept and growth parameters.

DISCUSSION

This article discusses ways to model the effect of time-invariant covariates in LGC modeling, and it introduces the *assumption of full mediation*. The standard way of modeling the time-invariant covariates (i.e., by using the growth predictor model in which an effect of the time-invariant covariate on the latent growth parameters is specified) has some attractive properties but may not always be appropriate. It is proposed that a less restrictive model with a direct effect of the time-invariant covariate on the indicators could be a better choice in such instances. This, so-called direct effect model also enables an explicit test of the assumption of full mediation, and it could be a viable alternative if the growth predictor model is rejected. Comparable models have already been proposed in SEM in a different context. For instance, in the context of research on method variance, where it has been attempted to measure the variables associated with method effects and to regress the other (substantive) indicators directly on this latent method factor (Williams & Anderson, 1994). In the area of LGC modeling this approach has received no attention, however.

Basically, the difference between the growth predictor model and the direct effect model lies in the (implicit) definition of the third variable (i.e., the time-invariant covariate), and in the substantive goal of the LGC analysis. In the direct effect

model, where the indicators are directly regressed on the time-invariant covariate, it should be regarded as a true covariate. The main interest here is in estimates of the LGC model corrected for nuisance, and the growth curves are estimated, corrected for the influence of the covariate. No special attention is given to the effect of the covariate; it is just used to get an unbiased estimate of the LGC model.

This situation is different, however, in the growth predictor model, where the effect of the time-invariant covariate is modeled on the latent growth parameters. In this model the time-invariant covariate is used as a predictor of the growth parameters, and the substantive interest is in both the LGC model and the prediction by the time-invariant covariate. In other words, the third variable is given an additional interpretation in this model. Besides the fact that the variable has covariation with the indicators, it is also regarded as having predictive value for the growth curves. This has been illustrated in the example: Whereas the growth predictor model represented the true model, the direct effect model also provided a good fit to the data. In summary, predictive value of a variable for the latent growth parameters in an LGC model implies covariation of this variable with the indicators; on the other hand, covariation with the observed indicators does not imply predictive value.

The model advocated here has not received much attention in LGC modeling. Neither has it in the cognate area of longitudinal multilevel analysis, where the implications are equivalent. However, the idea of regressing a covariate directly on the observed indicators is not new, but as noted by Rovine and Molenaar (2001), it is implicit in any standard ANCOVA. Moreover, Rovine and Molenaar stated that, "To include a covariate in fitting the SEM variants of multilevel models one must regress out the covariate not at the latent-variable level but at the level of the measurement model" (p. 90). A plausible explanation why this has not been picked up in practice might be that most theories explicitly posit predictive characteristics to covariates, leading credibly to the standard way of modeling the covariate as a predictor of the latent growth parameters. As we have shown, however, the standard approach (i.e., the growth predictor model) may be incorrect in certain instances, whereas the advocated direct effect model performs well. Furthermore, it is shown that the plausibility of the standard growth predictor model can be investigated by comparing it directly to the direct effect model using a likelihood ratio test. This comparison provides a test of the assumption that the effect of the covariate is fully mediated by the growth parameters.

REFERENCES

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147–158.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.

Goldstein, H. (1995). *Multilevel statistical models*. London: Arnold.

Hox, J. J. (2002). *Multilevel analysis. Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics*, *16*, 163–200.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561–614). New York: Plenum.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data* (pp. 69–108). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

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

Muthén, B., & Khoo, S. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, *10*(2), 73–102.

Rovine, M. J., & Molenaar, P. C. M. (2001). A structural equations modeling approach to the general linear mixed model. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 65–96). Washington, DC: American Psychological Association.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*, 363–381.

Williams, L. J., & Anderson, S. E. (1994). An alternative approach to method effects by using latent-variable models: Applications in organizational behavior research. *Journal of Applied Psychology*, *79*, 323–331.

APPENDIX A

The following pages show that the growth predictor model (Model 1) can be considered a special case of the direct effect model (Model 2). For ease of presentation, it is assumed that the means of the latent variables are equal to zero. To keep the presentation in line with the previously discussed models, we focus on a linear growth model with 4 time points. However generalizations to LGC models with a different number of time points are straightforward.

In the exposition following, the time-invariant covariate, z_i is treated as a perfectly measured η -variable (η_2), making it possible to specify the model entirely in the so-called Lisrel “all y” model. To compare these models, we first express Model 1 (the growth predictor model) and Model 2 (the direct effect model) in matrices.

In the case of 4 time points, Model 1 can be written as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ 0 \end{bmatrix} \tag{11}$$

where

$$\begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \beta_0 \\ 0 & 0 & \beta_1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} \beta_0 \eta_2 + \zeta_0 \\ \beta_1 \eta_2 + \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} \beta_0 \zeta_2 + \zeta_0 \\ \beta_1 \zeta_2 + \zeta_1 \\ \zeta_2 \end{bmatrix} \quad (12)$$

Combining 12 and 13 gives

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \zeta_2 + \zeta_0 \\ \beta_1 \zeta_2 + \zeta_1 \\ \zeta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ 0 \end{bmatrix} = \begin{bmatrix} \zeta_0 + 0 \cdot \zeta_1 + (\beta_0 + 0 \cdot \beta_1) \zeta_2 + \varepsilon_1 \\ \zeta_0 + 1 \cdot \zeta_1 + (\beta_0 + 1 \cdot \beta_1) \zeta_2 + \varepsilon_2 \\ \zeta_0 + 2 \cdot \zeta_1 + (\beta_0 + 2 \cdot \beta_1) \zeta_2 + \varepsilon_3 \\ \zeta_0 + 3 \cdot \zeta_1 + (\beta_0 + 3 \cdot \beta_1) \zeta_2 + \varepsilon_4 \\ \zeta_2 \end{bmatrix} \quad (13)$$

Model 2 is represented by

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \gamma_1 \\ 1 & 1 & \gamma_2 \\ 1 & 2 & \gamma_3 \\ 1 & 3 & \gamma_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ 0 \end{bmatrix} \quad (14)$$

where

$$\begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \end{bmatrix} \quad (15)$$

Combining 15 and 16 gives

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \gamma_1 \\ 1 & 1 & \gamma_2 \\ 1 & 2 & \gamma_3 \\ 1 & 3 & \gamma_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ 0 \end{bmatrix} = \begin{bmatrix} \zeta_0 + 0 \cdot \zeta_1 + \gamma_1 \zeta_2 + \varepsilon_1 \\ \zeta_0 + 1 \cdot \zeta_1 + \gamma_2 \zeta_2 + \varepsilon_2 \\ \zeta_0 + 2 \cdot \zeta_1 + \gamma_3 \zeta_2 + \varepsilon_3 \\ \zeta_0 + 3 \cdot \zeta_1 + \gamma_4 \zeta_2 + \varepsilon_4 \\ \zeta_2 \end{bmatrix} \quad (16)$$

implying that Model 2 is identical to Model 1 if the following restrictions are imposed:

$$\gamma_3 = 2\gamma_2 - \gamma_1 \quad (17)$$

$$\gamma_4 = 3\gamma_2 - 2\gamma_1 \quad (18)$$

Following the conditions for nesting (e.g., Bollen, 1989, p. 291), we may conclude that Model 1 is nested within Model 2. In other words, if Model 1 represents the true model, then Model 2 will give identical parameter estimates if Restrictions 18 and 19 are imposed. In addition, if Model 1 represents the true model, the effect of the covariate on, respectively, the level and shape factor is represented by

$$\beta_0 = \gamma_1 \tag{19}$$

$$\beta_1 = \gamma_2 - \gamma_1 \tag{20}$$

APPENDIX B
Covariance Matrix and Means Vector of the Data Generated With an Effect of the Covariate on the Latent Variable

	<i>z</i>	<i>y</i> ₁	<i>y</i> ₂	<i>y</i> ₃	<i>y</i> ₄	<i>M</i>
<i>z</i>	.997					0
<i>y</i> ₁	.998	2.057				10
<i>y</i> ₂	1.995	2.805	5.681			12
<i>y</i> ₃	2.995	3.804	8.069	12.585		14
<i>y</i> ₃	3.990	4.801	10.701	16.612	22.757	16

APPENDIX C
Covariance Matrix and Means Vector of the Data Generated With an Effect of the Covariate on the Observed Variables at the First Two Occasions

	<i>z</i>	<i>y</i> ₁	<i>y</i> ₂	<i>y</i> ₃	<i>y</i> ₄	<i>M</i>
<i>z</i>	.996					0
<i>y</i> ₁	.499	1.309				10
<i>y</i> ₂	.499	1.059	1.944			12
<i>y</i> ₃	.000	.811	2.084	3.610		14
<i>y</i> ₃	.000	.810	2.722	4.640	6.804	16