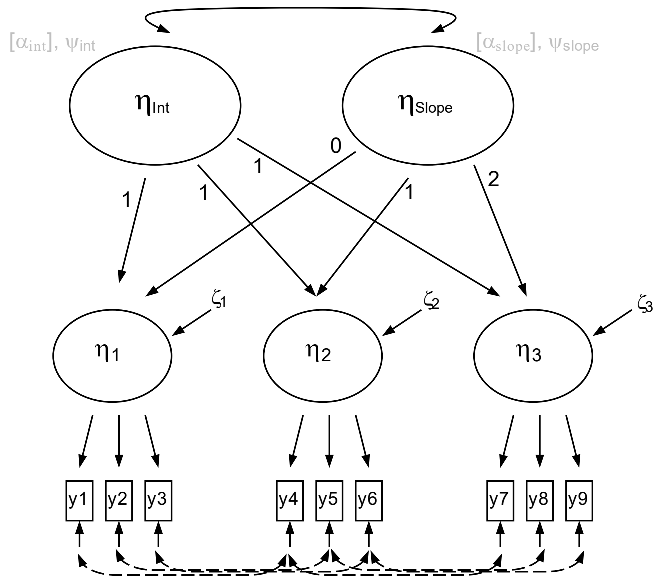
Second-Order Latent Growth Curve Models

Overview

Second-order growth curve models include multiple indicators of a latent variable at each time point (McArdle, 1988; Tisak & Meredith, 1990). They are sometimes referred to as *multiple indicator growth curve* models, *curve-of-factors* models, or *latent variable growth curve* models. The estimation of latent variables at each time point, which separately estimates measurement error (specific variance), will not impact mean slope estimates, because means are unbiased by measurement error. Having multiple indicators at each time point, however, will increase growth curve reliability and therefore increase statistical precision and increase power, (von Oertzen, Hertzog, Lindenberger, & Ghisletta, 2010; Wänström, 2009).

The second-order intercept and slope factors then model change in these variables and are specified much like they are with the more common first-order growth curve models, with intercept and slope factor means and variances estimated.

The first-order factors require attention to several specification details (see below). Scaling constraints can make a difference in the interpretation of the factor means so effect coding (Little, Slegers, & Card, 2006) is recommended so that the factor means at each time point are a weighted function of all of the indicators (Newsom, 2015). It is important to establish measurement invariance, including loadings and measurement intercepts (Ferrer, Balluerka, & Widaman, 2008).



Several variations on the basic specification are possible. Correlations among measurement residuals over time are recommended to improve the estimation of the intercept-slope covariance (Newsom, 2015). Constant variance across time points can be assumed by constraining the variances of the residual disturbances associated with the first order factors ($\zeta_1, \zeta_2, \zeta_3$) to be equal if desired and equivalence of disturbance variances can be tested using a likelihood ratio nested test.

Model Equations

An intercept and slope factor variance are specified for the second-order factors, with loadings set to 1.0 for the intercept factor, η_{int}, and slope factor, η_{slope}, loadings to unit increments (frequently, 0, 1, 2, 3…).

$$\eta_{ti} = \eta_{int} + \lambda_{1i}\eta_{slope} + \zeta_{ti}$$

Each factor estimates a mean, α, and deviation, ζ. α_{int} represents the average baseline value if time codes of 0, 1, 2, 3… are used for growth factor loadings, and α_{slope} represents the average slope.

$$\eta_{int} = \alpha_{int} + \zeta_{int}$$
$$\eta_{slope} = \alpha_{slope} + \zeta_{slope}$$

The first-order factors at each time point, *t*, are defined by measurement intercept values for each indicator, ν_{jt}, factor loadings, Λ_{jt}, a latent variable, η_{ti}, and indicator-specific variance, ε_{jti} where the subscript *j* refers to particular indicator of the first-order factor.

$$Y_{jti} = \nu_{jt} + \Lambda_{jt}\eta_{ti} + \varepsilon_{jti}$$

Model Specification Details

Identification. The most common method of identifying latent variables at each time point is to use a referent indicator for the first-order latent variable at each time point. Often the default in software programs, one loading, λ_{jti} = 1.0 and one measurement intercept, ν_{jti} = 0 for the first order factor are set (Bollen & Curran, 2006; Sayer & Cumsille, 2001), but the remaining intercepts can be freely estimated. The referent for the loading and the intercept should be the same and should be consistent across time points. The scaling of the factor mean at each time point is based on the observed mean of the referent measurement intercept, so the mean and variance of growth parameters will be a function of the observed mean of the chosen referent indicator variable.

An alternative identification approach that has been suggested is to standardize the factor mean and variance (α_1 = 0 and ψ_{11} = 1) at the first time point (Chan, 1988; Widaman & Reise, 1997). The advantage of this approach is that a single indicator does not have to be chosen as a referent arbitrarily, but the means represent a difference from the referent time point which may be less intuitive. Measurement invariance constraints are needed with this method.

The effects coding approach to identification of Little, Slegers, and Card (2006) has several advantages over the aforementioned approaches. This approach places constraints on the set of loadings and measurement intercepts for each factor, avoiding the arbitrary choice of one referent over another. The approach can be used in conjunction with longitudinal measurement constraints as well. An appealing advantage is that the means at each time point are defined as a weighted average of the indicators at each time point rather than based on a single referent indicator mean. It also allows the researcher to retain the original metric of the measured variables rather than scaling each time point as difference values from the first time point, as with the single occasion identification approach. Example code for a second-order growth model using Mplus and lavaan are posted online at www.longitudinalsem.com (example 7-6) illustrating each of these specifications (see ex7-6e.inp, ex7-6e.R for the effect coding).

Longitudinal measurement invariance. The second-order growth curve framework allows the researcher to examine factorial invariance over time (Bontempo & Hofer, 2007; Newsom, 2015), which is a key advantage. To the extent that measurement properties change over time, growth estimates are likely to be biased. As with between group factorial tests, one can distinguish between less and more strict invariance assumptions that include measurement intercepts, loadings, and measurement errors. Because mean estimates of latent slopes and intercepts are of interest and depend on values of measurement intercepts, invariance tests of measurement intercepts become more important than in models without interest in means (strict invariance). Because factor means are a weighted function of the loadings, loadings much also be invariant over time. If invariance is established, loadings and intercepts can be set equal and the research can proceed with estimating the growth model. When they are not invariant, the choice of referent (or, more generally, the factor scaling approach) may potentially affect the results and conclusions (Ferrer, Balluerka, & Widaman, 2008). If invariance fails, changes in the measure may be needed or the researcher may need to argue that the invariances effects are trivial and will lead to minimal bias.

Homogeneneous variance across time points. The researcher can set the disturbances associated with the first order factors ($\zeta_1, \zeta_2, \zeta_3$) to be equal over time to impose homogeneity of error variance, an assumption that can be tested for significance. Autocorrelation structures also can potentially be used by estimating correlations among the first-order factor disturbances across consecutive time points (lag 1) or other patterns, although these can potentially cause convergence issues, particularly with fewer time points.

Correlated measurement residuals. Correlated errors of same items over time are commonly used and are a reasonable addition to these models. Their inclusion should improve estimates of the variance and covariances of the growth factors.

References and Further Reading

Bollen, K.A., & Curran, P.J. (2006). *Latent Curve Models: A Structural Equation Approach.* New York: Wiley.
Bontempo, D. E., & Hofer, S. M. (2007). Assessing factorial invariance in cross-sectional and longitudinal studies. In A.D. Ong & M. van Dulmen (Eds.), *Handbook of methods in positive psychology* (pp. 153-175). Oxford University Press.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods, 1*, 421 -- 483.

Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. Methodology: European *Journal of Research Methods for the Behavioral and Social Sciences, 4*, 22–36.

Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling, 13*, 59-72.

McArdle, J.J. (1988). Dynamic but structural equation modeling of repeated measures data (pp. 561-614). In J.R. Nesselroade & R.B. Cattell (Eds.), *The Handbook of Multivariate Experimental Psychology*, Volume 2. New York, Plenum Press.

Newsom, J.T. (2015). Chapter 2 in *Longitudinal Structural Equation Modeling: A Comprehensive Introduction*. New York: Routledge.

Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp.179-200). Washington, DC: American Psychological Association.

Tisak, J, & Meredith, W. (1990). Descriptive and associative developmental models. In A. von Eye (Ed.), *Statistical methods in developmental research, volume 2* (pp. 387-406). San Diego, CA: Academic Press.

von Oertzen, T., Hertzog, C., Lindenberger, U., & Ghisletta, P. (2010). The effect of multiple indicators on the power to detect inter-individual differences in change. *British Journal of Mathematical and Statistical Psychology*, *63*, 627-646.

Wänström, L. (2009). Sample sizes for two-group second-order latent growth curve models. *Multivariate Behavioral Research, 44*, 588-619

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. *The science of prevention: Methodological advances from alcohol and substance abuse research*, 281-324.

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.