Structural Equation Modeling: A Multidisciplinary Journal

ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: https://www.tandfonline.com/loi/hsem20

A Short Note on Interpretation in the Dual Change Score Model

Ross Jacobucci, Sarfaraz Serang & Kevin J. Grimm

To cite this article: Ross Jacobucci, Sarfaraz Serang & Kevin J. Grimm (2019): A Short Note on Interpretation in the Dual Change Score Model, Structural Equation Modeling: A Multidisciplinary Journal, DOI: 10.1080/10705511.2019.1619457

To link to this article: https://doi.org/10.1080/10705511.2019.1619457

Published online: 16 Jul 2019.

Structural Equation Modeling: A Multidisciplinary Journal, 0: 1–7, 2019

ISSN: 1070-5511 print / 1532-8007 online
DOI: https://doi.org/10.1080/10705511.2019.1619457

A Short Note on Interpretation in the Dual Change Score Model

Ross Jacobucci,[1] Sarfaraz Serang,[2] and Kevin J. Grimm[3]
[1]*University of Notre Dame*
[2]*Utah State University*
[3]*Arizona State University*

The latent change score framework allows for estimating a variety of univariate trajectory models, such as the no change, linear change, exponential forms of change, as well as multivariate trajectory models that allow for coupling between two or more constructs. A particularly attractive feature of these models is that it is easy to decompose and interpret aspects of change. One particularly flexible model, the dual change score model, has two components of change: a proportional change component that depends on scores at the previous time point, and a constant change component that is additive. We demonstrate through simulation and an empirical example that in a correctly specified model, the correlation between the proportional change parameter and the mean of the constant change component can approach either −1 or 1, thus complicating interpretation. We provide recommendations and code to aid researchers' ability to diagnose this issue in their own data.

Keywords: longitudinal, latent change score, latent growth

The latent change score framework (McArdle, 2001; McArdle & Hamagami, 2001) has recently seen an increase in popularity for a number of reasons. One is that the models focus on individual change over time with strong interpretability of the components of change. To see how this manifests itself, we briefly review the model formulation.

As a basis, we decompose the observed score $Y[t]_i$ at time t for individual i into two components, such that

$$Y[t]_i = y[t]_i + u[t]_i \tag{1}$$

where $y[t]_i$ is a true score and $u[t]_i$ is a unique score (considered independent across time points). In this model, individual change over time can be seen as a difference between the true score at the current time (time t) minus the true score at the previous time (time $t-1$), such that

$$\Delta y[t]_i = y[t]_i - y[t-1]_i. \tag{2}$$

Across the entirety of the time span, total change manifests itself as the initial true level plus the summation of all previous true changes, such that

$$y[t]_i = y[0]_i + \sum_{r=1}^{r=t} \left(\Delta y[r]_i\right) \tag{3}$$

where $y[0]_i$ is the initial true level and $\sum_{r=1}^{r=t} \left(\Delta y[r]_i\right)$ is the summation of changes until the current time (time t).

With this as a basis, we can specify different functional forms of change by directly modeling $\Delta y[t]_i$. The three most commonly specified models, in order of complexity, are (a) the proportional change model, where $\Delta y[t]_i = \beta \cdot y[t-1]_i$ and β is the proportional change parameter; (b) constant change, where $\Delta y[t]_i = \alpha_{2i}$ and α_{2i} is the constant change latent variable (random effect); and (c) the dual change score model, which is a combination of both (a) and (b),

Correspondence should be addressed toRoss Jacobucci Department of Psychology, University of Notre Dame, 390 Corbett Family Hall, Notre Dame, IN 46556. E-mail: rcjacobuc@gmail.com

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hsem.

where $\Delta y[t]_i = \alpha_{2i} + \beta \cdot y[t-1]_i$. Across these three models, it is common for the dual change score model to fit best because it is the most flexible (equivalent to modeling exponential change; Serang, Grimm, & Zhang, 2018). Additionally, it is attractive from an interpretation perspective because change can be decomposed into two components that combine to result in a wide array of expected trajectories of change. However, in practice, we have noticed a limitation in the interpretation of the constant change component. The issue occurs when the correlation between the estimates for the mean of the constant change component and the proportional change parameter is close to either −1 or 1. When the absolute correlation is high, researchers should not interpret each component of change as independent and competing components that form the basis of change. In the following sections, we demonstrate this in an empirical longitudinal dataset, and we conduct a small-scale simulation study to identify what combination of a number of time points, true parameter estimates, and sample size lead to correlation among parameters.

EMPIRICAL EXAMPLE

We applied the dual change score model to longitudinal data collected on the Wechsler Intelligence Scale for Children dataset ($N = 204$; Osborne & Suddick, 1972), which contains four measurement occasions (ages 6, 7, 9, and 11 years). For this analysis, the outcome was the raw verbal scale score. Although most applications of the dual change score model have used frequentist estimation, we demonstrate the association between parameters using both frequentist and Bayesian estimation (see Hamagami, Zhang, & McArdle, 2009) because Bayesian estimation may highlight issues with convergence, or the effect of priors on the high correlation between parameters.

Frequentist Estimation

With the WISC data, we tested the proportional change, constant change, and dual change score models using maximum likelihood estimation in the (R Core Team, 2017) package lavaan (Rosseel, 2012). The only alteration to the models beyond the typical specification is the inclusion of two phantom latent variables (for ages 8 and 10) to account for the differing time lags between grades. Fit statistics are displayed in Table 1. Because the constant and proportional change models are not nested, we used three different information criteria to select the final model. Across all three information criteria, the dual change score model fit best. Parameter estimates from this model are displayed in Table 2, and from these estimates, the change equation can be written as $\Delta verbal[t]_i = 2.06 + .09 \cdot verbal[t-1]_i$.

Next, we examined the variance-covariance matrix of the parameter estimates. This is the sample estimate of the asymptotic covariance matrix (see pp. 468–469 in Bollen, 1989). We note that this is not an explicit parameter estimate in the LCS model, instead it represents the covariance of parameter estimates over repeated sampling. To make these values more interpretable, we converted the covariance matrix to a correlation matrix, thus allowing us to examine the correlations between all of the parameter estimates. The correlation between the mean of the constant change latent variable and the proportional change parameter was −0.989.

TABLE 1
Information Criteria Assess across Three Models on the WISC Data

	Proportional Change	*Constant Change*	*Dual Change*
Number of Parameters	4	6	7
−2LL	5049	5039	**5022**
BIC	5071	5071	**5059**
AIC	5057	5051	**5035**
aBIC	5058	5052	**5037**

Note. BIC is the Bayesian Information Criteria; AIC is the Akaike Information Criteria; aBIC is the sample size adjusted BIC.

TABLE 2
Parameter Estimates for the Dual Change Score Model

Parameters	*Maximum Likelihood Estimate*	*Standard Error*
α_1	20.34	0.39
α_2	2.06	0.66
β	0.09	0.02
σ_1^2	20.80	2.78
σ_2^2	0.83	0.21
σ_{12}	0.74	0.77
σ^2	12.18	0.85

Note. α is the mean of the intercept (α_1) and the slope (α_2); β is the proportional change parameter; σ is the variance of the intercept (σ_1^2), the slope (σ_2^2), and the covariance between slope and intercept (σ_{12}); σ^2 is the residual variance at each time point, constrained to be equal across time.

Bayesian Estimation

Using Stan (Stan Development Team, 2017) to fit the model with diffuse priors, we found the same high degree of correlation between parameter estimates in fitting the model across three chains. In contrast to the use of the variance-covariance matrix of parameter estimates in frequentist estimation, in Bayesian estimation this is equivalent to examining the correlation among samples taken for both parameters across the length of each chain (excluding burn-in). Across the three chains, the correlation between the mean of the constant change latent variable and the proportional change parameter across samples was −0.993. This can be seen in Figure 1, where the diagonal panes contain histograms of the sample estimates, whereas the off-diagonal panes are bivariate scatter plots of the samples for each parameter. Just as in frequentist

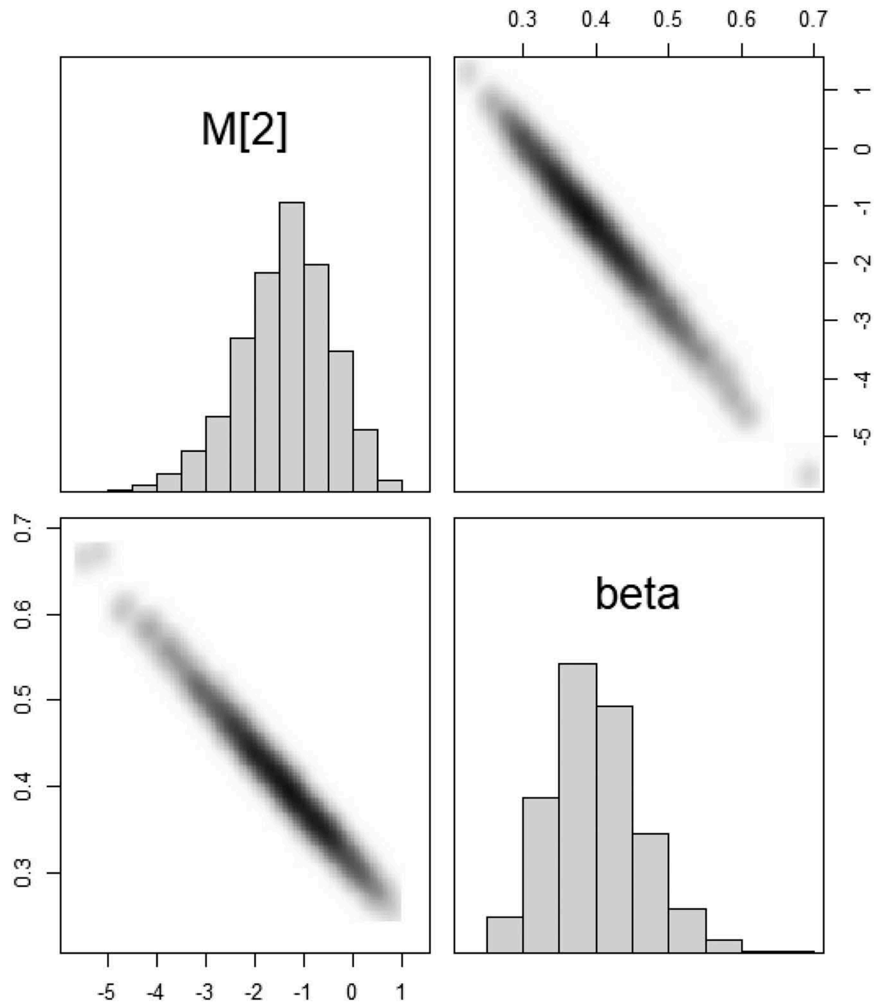

FIGURE 1 Univariate histograms and bivariate smoothed scatterplots from Stan for both the constant change mean (M[2]) and proportional change (beta) parameters. Note that phantom variables were not used in this specification, thus the parameter estimates differ from the frequentist estimates.

estimation, an extremely strong, negative correlation exists for the bivariate association between the mean of the constant change latent variable and the proportional change parameter. This poses problems[1] for the Hamiltonian Monte Carlo (Betancourt, 2017) that Stan uses by default. Since this sampler is informed by the gradient, it requires taking small step sizes to traverse the extremely narrow bivariate surface. Non-gradient informed samplers, such as the sampler used by JAGS (Plummer, 2003), do not have as much difficulty with this behavior in our experience.

Summary. Using both frequentist and Bayesian estimation, the correlation among the mean of the constant change latent variable and proportional change parameters was close to -1 when analyzing the WISC data. This can pose problems with estimation, but a larger issue is interpreting aspects of change. Practically speaking this means that these components of change are almost entirely dependent, meaning that researchers can only interpret one parameter in light of the other. In the WISC data, we interpret each parameter estimate in $\Delta verbal[t]_i = 2.06 + .09 \cdot verbal[t-1]_i$ as *conditional*, not purely additive, upon the value of the other component.

SIMULATION

Data Generation

To better understand whether the correlation between parameters in the WISC data was more the exception, rather than the rule, we examined this correlation across a variety of settings. These results highlight alternative research scenarios that may evidence the same problems as in the WISC.

[1] This can generally be overcome by changing the control arguments of max_treedepth to 20 and adapt_delta to 0.99. However, this results in slower sampling.

Although we did not have specific *a priori* hypotheses of the specific factors that would lead to strong correlations between parameters, we generally thought that fewer time points, and larger proportional change parameter estimates would exacerbate the problem. As a result, we used the WISC data as a template for the simulation conditions, with the goal of assessing more time points, smaller proportional change parameter estimates, and a range of estimates for the number of observations, sample size, and population values for the mean of the constant change latent variable. We expect the simulation conditions to be representative of the data used in practice.

Using the dual change score model as a template, we simulated data with 4, 6, 8, 10, 15, and 20-time points, of which observations were collected at 4, 6, 8, or 10 (always equal or less than the number of simulated) time points. Sample sizes were simulated to be 100, 200, 300, 500, 1000, and 5000, constant change means ranged from −5 to 5 in increments of 1, and proportional change parameters ranged from −.5, to .5 in increments of 0.1. Lastly, several parameters were not varied including the mean of the intercept ($\mu_0 = 1$), intercept variance ($\sigma_0^2 = 0.5$), variance of the constant change latent variable ($\sigma_1^2 = 0.5$), covariance between the intercept and the constant change latent variable ($\sigma_{10} = 0.1$), and residual variances ($\sigma_e^2 = 1$). Each condition combination was replicated 200 times.

RESULTS

Figure 2 is a plot of the mean correlations between the mean of the constant change latent variable and proportional change parameter against the population mean of the constant change latent variable. The values in the plot are distinguished by the number of time points and the population value of the proportional change parameter. In this plot, the true mean of the constant change latent variable is on the x-axis because this seemed to have a larger effect on the correlations as opposed to the true proportional change parameter. The most evident effect is that of the number of time points because the models with four-time points had correlations closest to either −1 or 1, and as the number of time points increased, there was an evident decrease in this association. Additionally, there did not seem to be specific combinations of the mean of the constant change latent variable and the proportional change parameter that resulted in high degrees of correlation. Instead, values of the mean of the constant change latent variable that were high in magnitude (both positive and negative) seemed to produce the strongest correlations.

The number of simulated time points displayed (Figure 3) had a more complex relationship with the correlation between the two change parameters. Most evident was the low correlations when the number of time points was large (15 or 20) and the population proportional change parameter was strongly positive. This may occur because for positive proportional

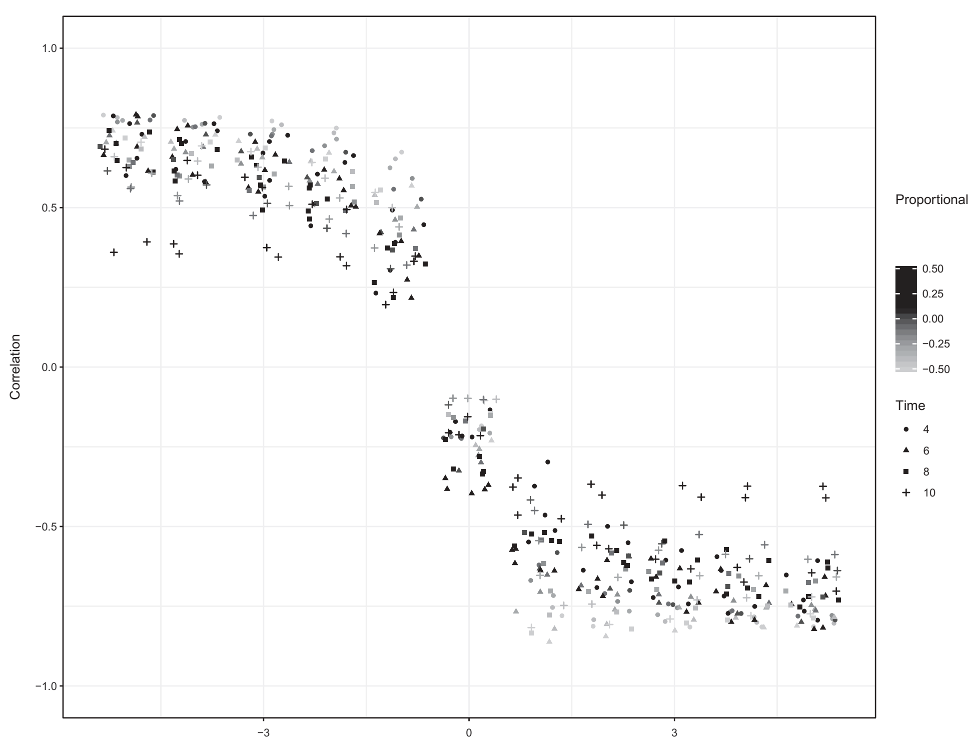

FIGURE 2 Average correlation across different simulation conditions. Note that points have been jittered to increase clarity.

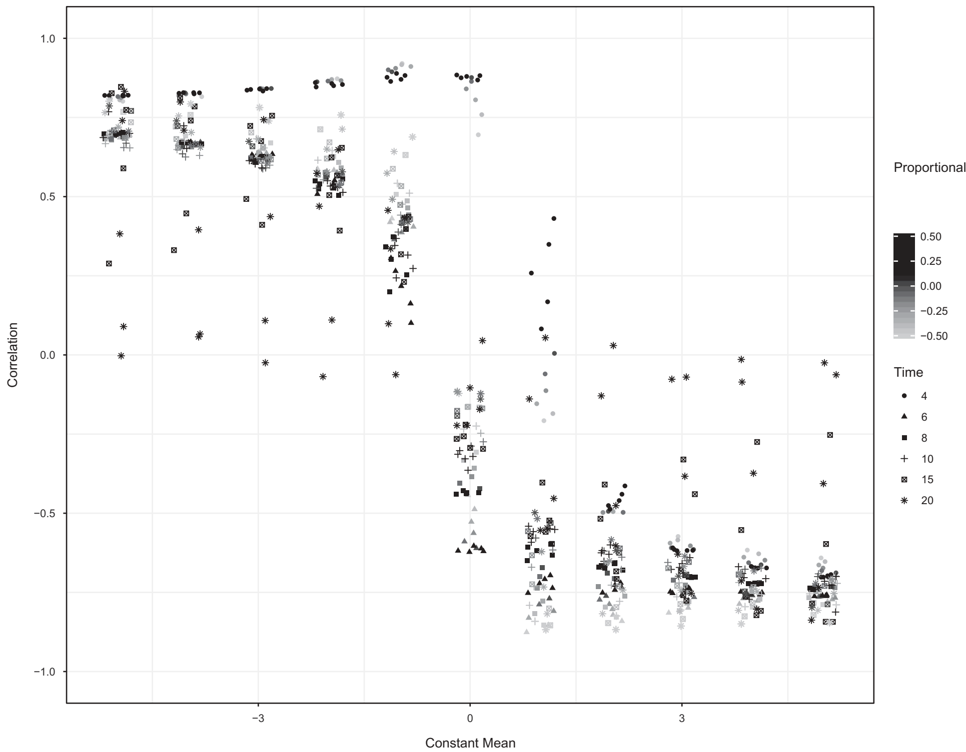

FIGURE 3 Average correlation across different simulations for each of the number of simulated time points. Only 4–10 time points were selected among the total observation time. Note that points have been jittered to increase clarity.

change parameters, the function is unbounded (approaches ± infinity) over time. This is regardless of whether the constant mean is positive or negative. With β = .5, it is possible that the proportional change component overtakes the mean component quite rapidly. That is, the expected scores are so large (in absolute value) that the constant change component loses its influence very early. Additionally, there was a slightly different pattern of correlations when there were four-time points. For simulated constant means of zero and one, there were positive correlations, differing from the behavior of more time points. In contrast to larger numbers of time points, the proportional change parameter exhibits less influence with a smaller number of time points, thus making it more difficult to differentiate from the constant change influence.

Lastly, we examined the correlations between various simulation conditions and the resultant parameter correlation. First, the number of time points had a correlation of −.24 with the absolute value of the parameter correlation confirming the problems with fewer time points, while number of simulated time points had a correlation of −0.02. However, the largest influence was the mean of the constant change component because it had a correlation of −.88 with the absolute value of the parameter correlation. This is extremely large, particularly given the nonlinearity of the trajectory displayed in Figure 2. Next, the absolute value of the proportional change parameter had a small correlation, 0.10, with the absolute value of the parameter correlation. Finally, sample size had a near zero correlation (r = 0.01).

DISCUSSION

The purpose of this paper was to highlight a limitation of the dual change score model that occurs in some analysis scenarios. Across both frequentist and Bayesian estimation, we found high correlations between both components of change – the mean of the constant change latent variable and the proportional change parameter. These additive components can be difficult to untangle when the number of time points is small, and when the mean of the constant change component is further away from zero.

One area of the expected impact of the correlation between parameters in the dual LCS model is fit indices. Because the only additional parameter in extending the constant change model to the dual change score model has a high correlation with a parameter already in the model, it is expected to influence the fit negligibly. In essence, the additional parameter is not completely independent, despite it counting as one against the degrees of freedom. This is similar in a sense to how parameters are counted in a Bayesian context, where priors can constrain the influence of individual parameters (e.g., Asparouhov, Muthén, & Morin, 2015). Given this, it is surprising that the dual change score model often demonstrates the best fit, owing to its ability to model a combination of both linear and proportional growth beyond less complex models. However, demonstrating best fit is not the only qualification in model selection; model fit should be paired with substantive interpretation. It is possible that fitting a sequence of more complex models than the dual change score model would result in a different conclusion with respect to a best fitting model.

Our goal in highlighting this is not to prevent researchers from using the model. Instead, we advocate for assessing the correlation between parameters. In frequentist estimation, this involves extracting the variance-covariance matrix of the parameter estimates. Most structural equation modeling software packages report this information (see supplementary lavaan code). In Bayesian estimation, it is important to check bivariate scatterplots to look for a near dependence across samples for each parameter.

One cause of this problem may be empirical underidentification (e.g., Green & Yang, 2017). For example, consider the problem of a two-factor model with two indicators per factor and no cross-loadings or covariances among measurement errors (Bollen, 1989). If the factors are uncorrelated, this model has two degrees of freedom. However, unique parameter estimates cannot be obtained for this model, making it empirically underidentified. This phenomenon may be at play in the dual change score model. The trajectory of the observed scores is the sum of the proportional and constant change components. This creates a form of soft constraint in that if the proportional change parameter increases, the constant change component must decrease to compensate. This constraint may be analogous to the constraint imposed by fixing the correlation between factors in the two-factor example to zero. Despite the dual change score model with four time points having seven degrees of freedom, given some trajectories of change, more time points appear to be necessary in order to isolate the contribution of the constant change component and proportional change parameter. This points to issues that arise when there is a mismatch between model complexity and the amount of information in the data, despite correct model specification.

Additionally, this lack of identification may be a contributing factor to convergence issues (see O'Rourke, Grimm, & MacKinnon, 2017) in more complex LCS models, both univariate and bivariate. Another perspective on these issues is that of fungible parameter estimates (Lee, MacCallum, & Browne, 2018), which occurs when changing a parameter estimate from the optimal value to another results in a very small change in the log-likelihood. A potential remedy is to penalize (regularize) both components in order to reduce either or both of their influence. Similar to the proposed methods in Jacobucci and Grimm (2018), this may reduce the correlation and allow for clearer interpretation of each components contribution.

Importantly, we note that having a strong correlation between these two parameters of the dual change score model does not necessarily mean that the model is inappropriate for the data. Instead, it should cause caution in interpreting the components of change as independent contributions. In our experience, we have not found a reparameterization, or alternative specification, that can overcome this correlation between parameter estimates. Additionally, in the bivariate LCS framework, the coupling parameters can be seen as proportional change parameters, albeit depending on a different variable. Given this, we expect the same problem (most likely exacerbated given the larger number of parameters) to plague the bivariate dual change score models. Finally, we expect that this problem of strong correlations between parameters is not unique to the LCS model. We expect the same issues to appear in alternative scenarios where limitations to the data exist such as a ceiling or floor effects, a high degree of missing data, or a lack of between person variability in specific variables.

Our simulation identifies specific data characteristics as being particularly susceptible to high absolute correlations among parameters of change. Most notable was the mean of the constant change latent variable. When the mean was relatively high, the correlation among the proportional and constant change parameters was highly negative, with large positive correlations when the constant change latent variable mean was negative. Additionally, the number of collected time points influenced the correlation, with four-time points having the highest absolute correlations. One buffer to this correlation was the duration of the study, as more time passing lessened the correlation between parameters. Given that the LCS model is a model for studying the *dynamics* of change (e.g., Ferrer & McArdle, 2010), it makes sense that the more time that passes better allows the components of change to differentiate.

In conclusion, we advocate for a measure of caution in interpreting specific parameters in the dual change score model. This speaks to a larger aspect of data analysis – going beyond just getting a model to fit and more thoroughly assessing assumptions and nuanced aspects of the fit. This is easier to do and a more common practice in Bayesian estimation, as graphical displays of sampling are easy to produce and necessary to identify convergence. In frequentist estimation, much less information is readily available, requiring additional assessment into whether assumption violations occurred. We hope that the problems briefly outlined in this paper provide more incentive for researchers to dig a little deeper to understanding these aspects of their models.

ORCID

Ross Jacobucci http://orcid.org/0000-0001-7818-7424
Kevin J. Grimm http://orcid.org/0000-0002-8576-4469

REFERENCES

Asparouhov, T., Muthén, B., & Morin, A. J. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances:

Comments on Stromeyer et al. *Journal of Management*, *41*, 1561–1577. doi:10.1177/0149206315591075

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*.

Bollen, K. A. (1989). *Wiley series in probability and mathematical statistics. Applied probability and statistics section. Structural equations with latent variables*. Oxford, England: John Wiley & Sons.

Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, *19*, 149–154. doi:10.1177/0963721410370300

Green, S., & Yang, Y. (2017). Empirical underidentification with the bifactor model: A case study. *Educational and Psychological Measurement*, 1–20. doi:10.1177/0013164417719947

Hamagami, F., Zhang, Z., & McArdle, J. J. (2009). Modeling latent difference score models using Bayesian algorithms. In S. M. Chow, E. Ferrer, & F. Hsieh (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 319–348). New Jersey: Lawrence Erlbaum Associates.

Jacobucci, R., & Grimm, K. J. (2018). Regularized estimation of multivariate latent change score models. In E. Ferrer, S. Boker, & K. J. Grimm (Eds.), *Advances in longitudinal models for multivariate psychology: A Festschrift for Jack McArdle*(pp. 109–125). London, UK: Routledge.

Lee, T., MacCallum, R. C., & Browne, M. W. (2018). Fungible parameter estimates in structural equation modeling. *Psychological Methods*, *23*, 58. doi:10.1037/met0000130

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structure analysis. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future. A festschrift in honor of Karl Jöreskog* (pp. 1–40). Lincolnwood, IL: Scientific Software International, Inc.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). Washington, DC: American Psychological Association.

O'Rourke, H., Grimm, K. J., & MacKinnon, D. P. (2017). *Time metrics and their effects in latent change score models*. Unpublished manuscript.

Osborne, R. T., & Suddick, D. E. (1972). A longitudinal investigation of the intellectual differentiation hypothesis. *Journal of Genetic Psychology*, *121*, 83–89.

Plummer, M. (2003). Jags: A program for analysis of Bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125).

R Core Team. (2017). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from http://www.R-project.org/

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*, 1–36. doi:10.18637/jss.v048.i02

Serang, S., Grimm, K. J., & Zhang, Z. (2018). On the correspondence between the latent growth curve and latent change score models. *Structural Equation Modeling*, doi:10.1080/10705511.2018.1533835

Stan Development Team. 2017. Stan modeling language users guide and reference manual, Version 2.17.0. http://mc-stan.org

APPENDIX A

Lavaan Code

```
   mod.dual.phantom <- "
   #latent variables
   lV1 =~ 1*V1
   lV2 =~ 1*V2
   lV3 =~ 0
   lV4 =~ 1*V4
   lV5 =~ 0
   lV6 =~ 1*V6
   #autoregressions
   lV2 ~ 1*lV1;lV3~1*lV2; lV4 ~ 1*lV3;lV5 ~ 1*lV4; lV6
~ 1*lV5
   #change - delta; d
   dV1 =~ 1*lV2; dV2 =~ 1*lV3; dV3 =~ 1*lV4;dV4 =~
1*lV5; dV5 =~ 1*lV6
   #intercept and slope
   inV =~ 1*lV1;
   # match lgm
   slope =~ 1*dV1 + 1*dV2 + 1*dV3 + 1*dV4 + 1*dV5
   #manifest means @0
   V1 ~ 0*1; V2 ~0*1; V4 ~ 0*1; V6 ~ 0*1
   #slope and intercept means
   slope ~ 1;
   inV ~ 1;
   #Latent variances and covariance
   slope ~~ slope;
   inV ~~ inV;
   slope ~~ inV;
   #means and vars @0
   lV1 ~ 0*1; lV2 ~0*1; lV3 ~ 0*1; lV4 ~ 0*1;lV5 ~ 0*1;
lV6 ~ 0*1
   dV1 ~ 0*1; dV2 ~0*1; dV3 ~ 0*1;dV4 ~ 0*1; dV5 ~0*1;
   lV1 ~~ 0*lV1; lV2 ~~ 0*lV2; lV3 ~~ 0*lV3; lV4 ~~
0*lV4;lV5 ~~ 0*lV5; lV6 ~~ 0*lV6
   dV1 ~~ 0*dV1; dV2 ~~ 0*dV2; dV3 ~~ 0*dV3;dV4 ~~
0*dV4; dV5 ~~ 0*dV5
   #auto-proportions
   dV1 ~ beta*lV1
   dV2 ~ beta*lV2
   dV3 ~ beta*lV3
   dV4 ~ beta*lV4
   dV5 ~ beta*lV5
   #residuals equal
   V1 ~~ resid*V1; V2 ~~ resid*V2; V4 ~~ resid*V4; V6 ~~
resid*V6;"fit.dual.phantom <- lavaan(mod.dual.phan-
tom, data=wisc)
   summary(fit.dual.phantom,standardized=TRUE,
fit=TRUE)
   vcov.dual.phantom= inspect(fit.dual.
phantom,"vcov")
   cov2cor(vcov.dual.phantom)
```