

# On the Mathematical Relationship Between Latent Change Score and Autoregressive Cross-Lagged Factor Approaches: Cautions for Inferring Causal Relationship Between Variables

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. <<https://doi.org/10.1080/00273171.2015.1079696>>

**On the Mathematical Relationship Between Latent Change Score and Autoregressive Cross-Lagged Factor Approaches: Cautions for Inferring Causal Relationship Between Variables**

Causal Inference

Models for causal inference

LCS

Relationship between LCS and cross-lagged

Empirical Analysis

Discussion

## Causal Inference

philosopher David Hume summarized three conditions for identifying a causal relationship from A to B:

- (1) A precedes B in time (temporal precedence),
- (2) A and B are contiguous in space and time (contiguity),
- (3) A and B always co-occur or neither of them occurs (covariation; see Holland).

Cook and Campbell (1979) also suggested three similar conditions of causality

- (1) temporal precedence
- (2) covariation between A and B
- (3) the exclusion of alternative explanations for cause-effect linkages.

Rubin (1974) extended a counterfactual model from Jerzy Neyman's work (Neyman, 1990) and proposed the potential outcomes model (alternatively referred to as the Neyman-Rubin-Holland model)

Granger causality, which originated in economics, emphasizes temporal precedence, unique effect, and inertia (or stability).

## Models for causal inference

tested via the statistical significance of the causal path from A at timepoint  $t - 1$  to B at timepoint  $t$  while controlling for B at timepoint  $t - 1$ .

- Latent change score: extension of the latent growth curve model
- Cross-lagged model: extension of vector autoregressive model

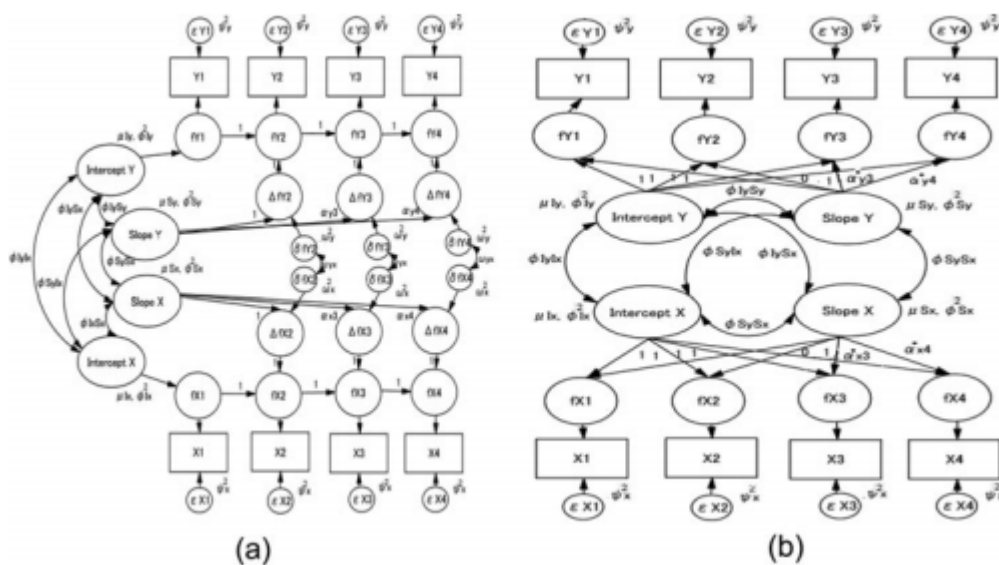
Unlike the LCS model, which incorporates information from both means and covariances, the ARCL model has been commonly used in longitudinal data analysis to examine change based on manifest variable covariances only

McArdle (2009) noted that "This new change model [the bivariate LCS model] subsumes all aspects of the previous models [including the ARCL model] as special cases to be tested. We can fit the standard crosslagged factor model ... by eliminating the latent intercepts and the latent slopes" (p. 597).

both the LCS model and the ARCL model are subsumed into the autoregressive latent trajectory (ALT) model

## LCS

Relation between LCS and latent growth curve model:



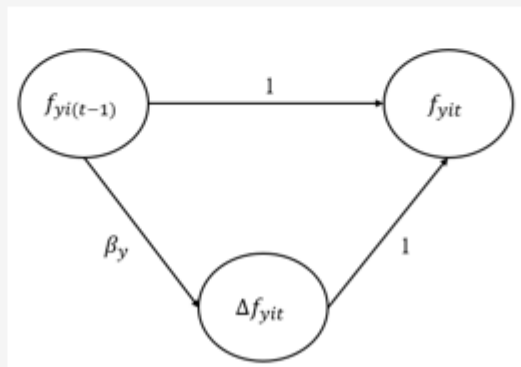
Equation

Let  $y_{it}$  and  $x_{it}$  be observed variables at timepoint  $t$  ( $1, \dots, t, \dots, T$ ) for each individual  $i$  ( $1, \dots, i, \dots, N$ )

$$y_{it} = f_{yit} + \epsilon_{yit}, x_{it} = f_{xit} + \epsilon_{xit}$$

For  $t \geq 2$ ,  $f_{yit}$  and  $f_{xit}$  are further decomposed into latent change score  $\Delta f_{yit}$  and  $\Delta f_{xit}$  that express increments between timepoint  $t - 1$  and  $t$

$$f_{yit} = f_{yi(t-1)} + \Delta f_{yit}, f_{xit} = f_{xi(t-1)} + \Delta f_{xit}$$

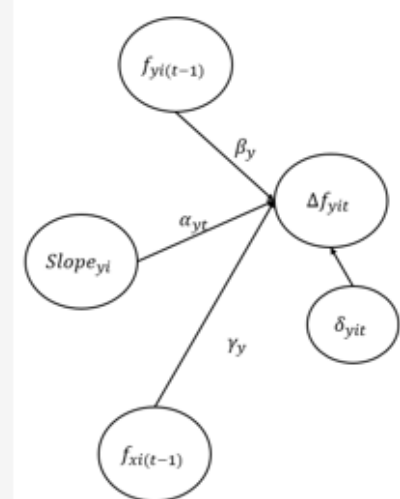


Latent change scores  $\Delta f_{yit}$  and  $\Delta f_{xit}$  are modeled by some specified (linear) functional form. The most popular form in actual practice is

$$\Delta f_{yit} = \alpha_{yt} S_{yi} + \beta_y f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + \delta_{yit}$$

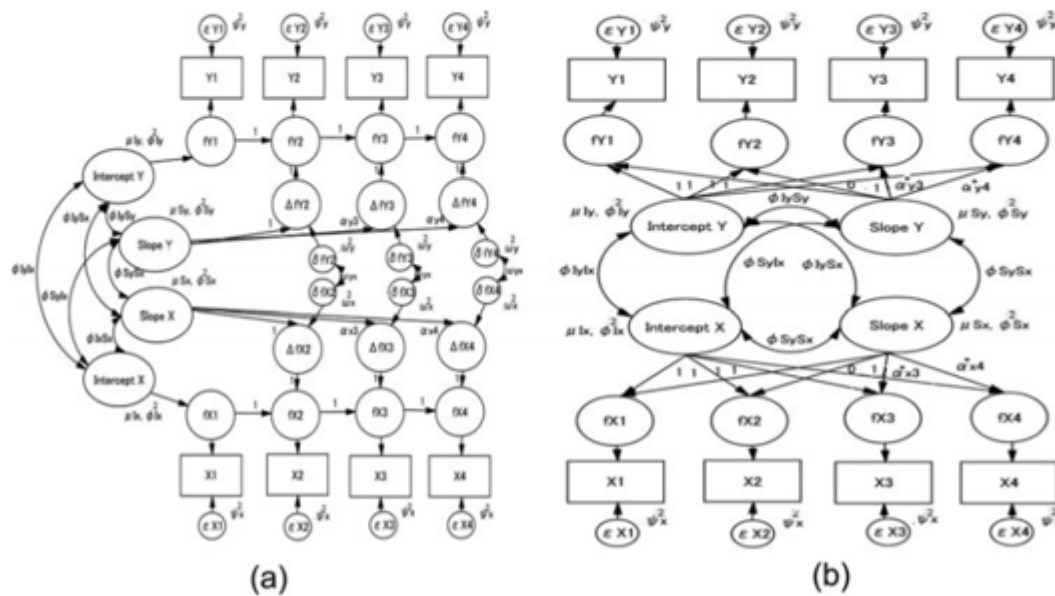
$$\Delta f_{xit} = \alpha_{xt} S_{xi} + \beta_x f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + \delta_{xit}$$

- $\alpha_{yt}$  and  $\alpha_{xt}$  are factor loadings from slope factors  
 $t = 1, \alpha = 0; t = 2, \alpha = 1$
- $\beta_y$  and  $\beta_x$  are autoregressive parameters
- $\gamma_y$  and  $\gamma_x$  are coupling parameters used to evaluate dynamic changes between variables
- time-variant vs time-invariant (McArdle, 2009)



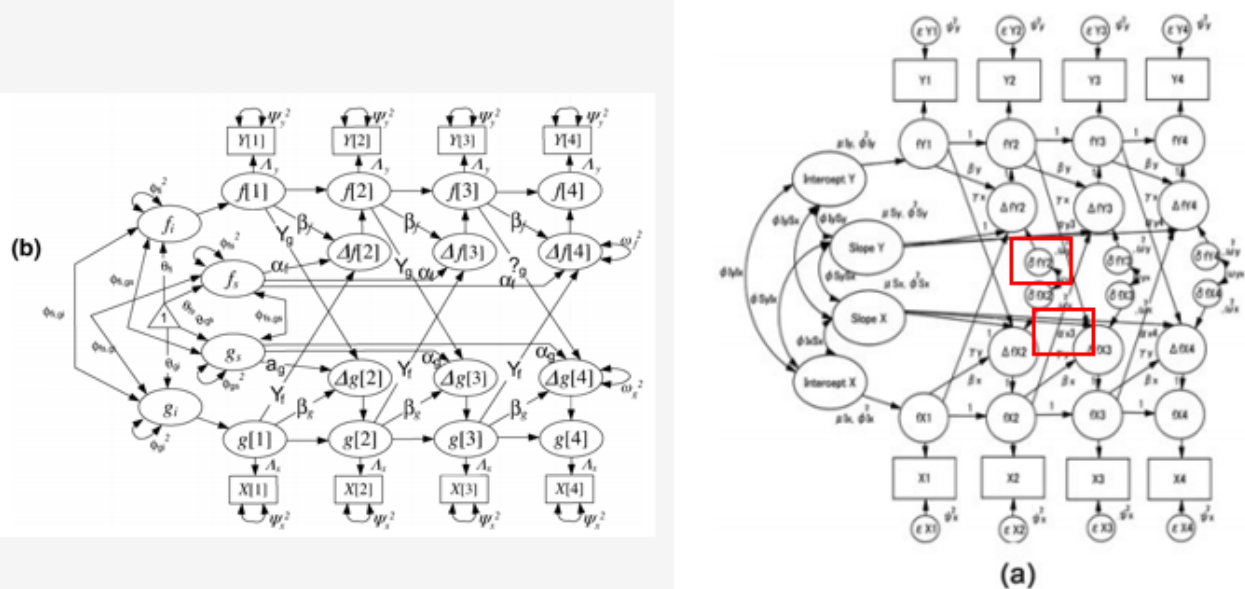
## Combination of latent growth curve and auto-regressive models

- Auto-regressive: examine change based on covariances
- Latent growth curve: examine growth (i.e., means)  $\alpha_{yt}^* = \sum_{t'=1}^t \alpha_{yt'}$
- autoregressive parameters, coupling parameters, and residual = 0



Constraints in practice:

no factor disturbance and time-invariant parameters



# Relationship between LCS and cross-lagged

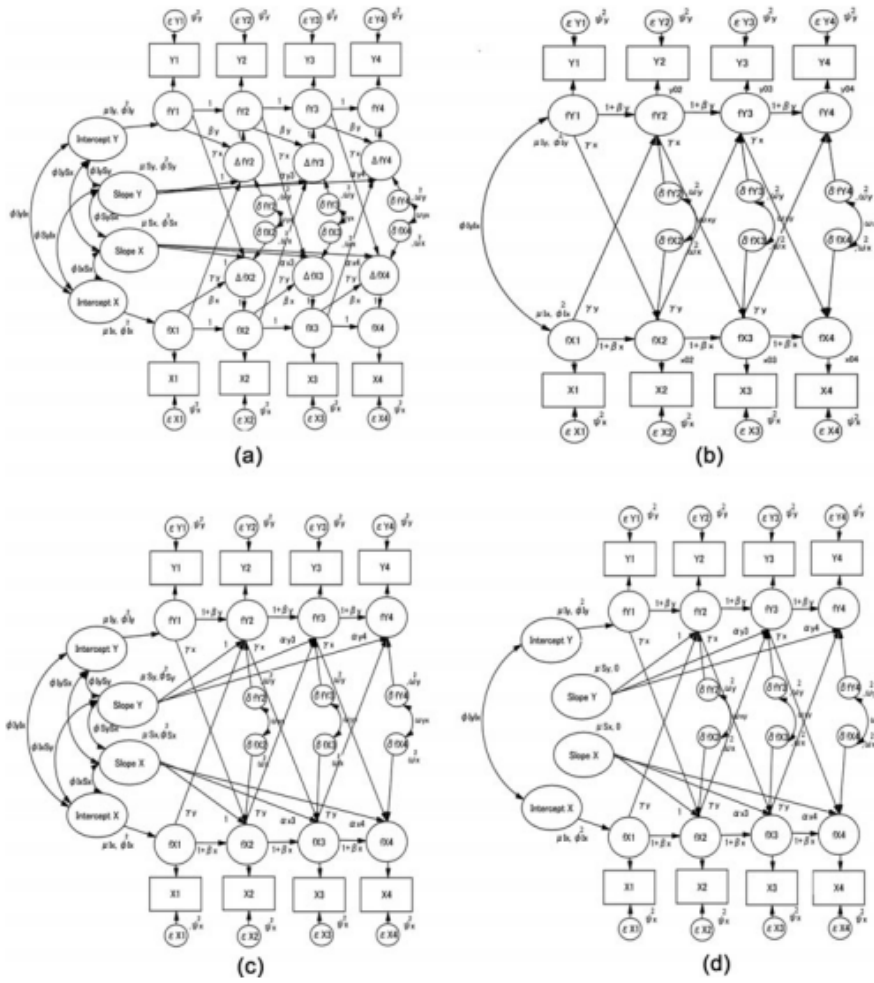


FIGURE 2 (a) Latent change score (LCS) model (Eq. 2 and 3). (b) Autoregressive cross-lagged (ARCL) factor model (Eq. 4). (c) LCS model (Eq. 5). (d) ARCL model based on LCS model (c) specification. ARCL model in (d), which shows mathematical equivalence to the ARCL model under standard specification (b) (setting  $f_{Y11} = I_{Y1}$ ,  $f_{X11} = I_{X1}$ ,  $\gamma_{Y2} = S_{Y2}$ ,  $\gamma_{Y3} = \alpha_{Y3}\gamma_{Y2}$ ,  $\gamma_{Y4} = \alpha_{Y4}\gamma_{Y3}$ ,  $\gamma_{X2} = S_{X2}$ ,  $\gamma_{X3} = \alpha_{X3}\gamma_{X2}$ ,  $\gamma_{X4} = \alpha_{X4}\gamma_{X3}$ ), is a special case of the LCS model (c) (setting five factor variances and two factor covariances to 0), which again shows mathematical equivalence to the LCS model under standard specification (a).

## LCS

$$f_{yit} = f_{yi(t-1)} + \Delta f_{yit}, f_{xit} = f_{xi(t-1)} + \Delta f_{xit}$$

$$\Delta f_{yit} = \alpha_{yt} S_{yi} + \beta_y f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + \delta_{yit}$$

$$\Delta f_{xit} = \alpha_{xt} S_{xi} + \beta_x f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + \delta_{xit}$$

Combine these equations

$$f_{yit} = \alpha_{yt} S_{yi} + (1 + \beta_y) f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + \delta_{yit}$$

$$f_{xit} = \alpha_{xt} S_{xi} + (1 + \beta_x) f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + \delta_{xit}$$

## ARCL

$$f_{yit} = y_{0t} + \beta_y f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + \delta_{yit}$$

$$f_{xit} = x_{0t} + \beta_x f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + \delta_{xit}$$

The LCS model is reduced to the ARCL model when

$$\alpha_{yt} S_{yi} = y_{0t}, \alpha_{xt} S_{xi} = x_{0t}$$

The LCS model is reduced to the ARCL model when

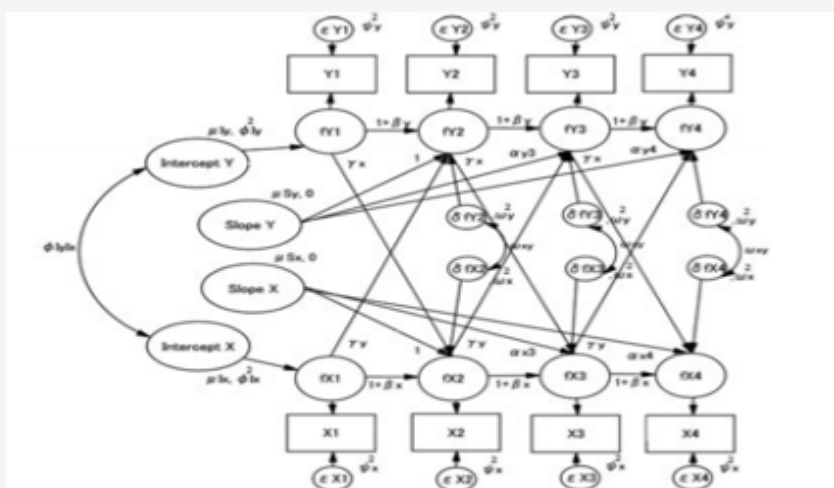
$$\alpha_{yt} S_{yi} = y_{0t}, \alpha_{xt} S_{xi} = x_{0t}$$

The slope factor scores are assumed to be invariant among people

$$\phi_{S_y}^2 = \phi_{S_x}^2 = \phi_{S_y S_x} = \phi_{S_y I_y} = \phi_{S_x I_x} = \phi_{I_y S_x} = \phi_{S_y I_x} = 0$$

$$S_{yi} = y_{02}, S_{xi} = x_{02}, \alpha_{yt} = \frac{y_{0t}}{y_{02}}, \alpha_{xt} = \frac{x_{0t}}{x_{02}}$$

Difference in the degrees of freedom: 7



Difference:

The random effect assumption regarding the slope factor scores in the LCS model.

Interpretation: the expected latent change scores (or increments between timepoints  $t - 1$  and  $t$ ) are the same among individuals who share not only the same latent true scores  $f(y_i(t-1))$  and  $f(x_i(t-1))$  but also the same slope factor scores.

Similarity: evaluate the causal precedence between variables, and both models have been commonly used for this purpose in longitudinal designs.

Interpretation of the random effect

- From a standard interpretation, random effects depict moderators that are not included in the model.
- an alternative interpretation of random effects on slope factor scores is the perturbation of error variances. As one reviewer commented, one taking a systems approach assumes perturbation and is unlikely to be interested in making a causal argument in the first place, which is different from standard behavioral research for causal inference.

Model selection: next paper

due to a lack of prior empirical investigation in this area little is known about (a) what the best procedure is for model selection in this context, and (b) when and how often such discrepant results can occur. In forthcoming work (Usami, Hayes, & McArdle, in press)

## Empirical Analysis

These examples illustrate how conclusions regarding causal precedence can differ dramatically between the LCS and ARCL models

**Difference in  $\gamma_y$  and  $\gamma_x$  across ARCL and LCS models**

It is expected that high correlations between  $S_{yi}$  ( $S_{xi}$ ) and  $f(y_i(t-1))$  ( $f(x_i(t-1))$ ) can lead indirectly to unstable estimates of  $\beta_y$  ( $\beta_x$ ) resulting from multicollinearity.

The standard errors of  $\beta_y$  ( $\beta_x$ ) in the LCS model are about 10-20 times larger than those in the ARCL models

## Discussion

Recent work by Hamaker et al. (2015) insisted that cross-lagged models, which do not represent the actual within-person relationships over time, may lead to erroneous conclusions regarding the presence, predominance, and sign of causal influences

- lcs may be an alternative

■

[Return](#)