A Unified Framework of Cross-Lagged Longitudinal Models

Satoshi Usami (University of Tokyo) Kou Murayama (University of Reading)
Ellen Hamaker (Utrecht University)

1 Introduction

One of the primary interests in longitudinal data analysis is the inference about the reciprocal effects or causality between variables. To address the reciprocal effects, cross-lagged panel model (CLPM) has been widely used. However, other recent longitudinal models incorporate cross-lagged relations, such as latent change score (LCS) model (McArdle & Hamagami, 2001; Usami, Hayes, & McArdle, 2015), autoregressive latent trajectory (ALT) model (Curran & Bollen, 2001), and (a bivariate version of) stable trait autoregressive trait and state (STARTS) model (Kenny & Zautra, 2001). In addition, a recent study by Hamaker, Kruiper, & Grasman (2015) criticized CLPM that it does not represent the actual within-person causal relations over time, potentially leading to erroneous conclusions regarding reciprocal effects. They instead proposed an alternative model, random-intercepts CLPM (RI-CLPM). The central aim of this presentation is to provide a unified statistical framework that clarifies the mathematical and conceptual relationships among those models.

2 Unified Framework

Assume we are interested in the reciprocal effects between variables X and Y, and let x_{it} and y_{it} be repeated measures at time t for each individual i. In the unified framework x_{it} and y_{it} for $t \geq 2$ can be expressed as:

$$x_{it} = f_{xit} + c_1\epsilon_{xit}, \quad y_{it} = f_{yit} + c_1\epsilon_{yit}, \tag{1}$$

$$f_{xit} = [c_2(\mu_{xt} + \alpha_{xt}T_{xi}) + c_3 r_{xit}] + f^*_{xjt}, \quad f_{yit} = [c_2(\mu_{yt} + \alpha_{yt}T_{yi}) + c_3 r_{yit}] + f^*_{yjt}, \tag{2}$$

$$f^*_{xjt} = [(1-c_2)(I^*_{xi} + \alpha_{xt}S^*_{xi}) + (1-c_3)r^*_{xit}] + \beta_x f^*_{xj(t-1)} + \gamma_x f^*_{yj(t-1)},$$
$$f^*_{yjt} = [(1-c_2)(I^*_{yi} + \alpha_{yt}S^*_{yi}) + (1-c_3)r^*_{yit}] + \beta_y f^*_{yj(t-1)} + \gamma_y f^*_{xj(t-1)} \tag{3}$$

First part, which is shown in Equation 1, is a *measurement part*. Namely, observations are first decomposed into latent true scores (f_{xit} and f_{yit}) and errors (ϵ_{xit} and ϵ_{yit}). c_1 is a dummy variable that switches on/off the measurement error of the model. Second part, which is shown in Equation 2, is a *decomposition part* mainly to define trait factors or growth factors incorporated in the model. μ_{xt} and μ_{yt} are the temporal group means at each time point, and T_{xi} and T_{yi} are trait factors that express individual's stable trait-like deviations from these means. Trait factors scores $(T_{xi}, T_{yi})^t$ have a 2×1 mean vector $\mathbf{0}$ and 2×2 variance-covariance matrix $\boldsymbol{T}$, and these trait factors have factor loadings or weights α_{xt} and α_{yt}. Importantly, this part also has residuals, r_{xit} and r_{yit}, at time point t. We shall call these residuals non-dynamic residuals. Third part (Equation 3) defined for $t \geq 2$ is a *dynamic part*. Namely, all terms included in this equation influence forwardly the true scores at the later time points through the autoregressive and lagged relationships as dynamic processes. In this part, growth factors I^* and S^* with non-zero means are used instead of trait factors to model latent trajectories. I^*_{xi} and I^*_{yi} are intercept factors, and S^*_{xi} and S^*_{yi} are slope factors. β_x and β_y are autoregressive parameters, and γ_x and γ_y are cross-lagged parameters. c_2 and c_3, which appear in both decomposition part and dynamic part, are dummy variables that define trait factors or growth factors and residuals related to these factors. A model includes trait factors when $c_2 = 1$, whereas it includes growth factors when $c_2 = 0$. On the other hand, a model has (non-dynamic) residuals when $c_3 = 1$ whereas it does not have non-dynamic residuals when $c_3 = 0$.

For example, setting $c_1 = 0$, $c_2 = 1$ and $c_3 = 0$ and constraining $\alpha_{xt} = \alpha_{yt} = 1$ is algebraically equivalent to RI-CLPM. This model can be further reduced to CLPM by setting $\alpha_{xt} = \alpha_{yt} = 0$. In the presentation, we will explain the differences of the ways of interpretation of the cross-lagged effects among models, and also provide a simulation study that shows existing longitudinal cross-lagged models are generally susceptible to the issue of improper solutions, even if a model is correctly specified, but we can reduce the risk of improper solutions by using non-dynamic instead of dynamic residuals.

References

Curran, P.J., & Bollen, K.A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L.M. Collins & A.G. Sayer (Eds.), *New methods for the analysis of change*, (pp. 105-136). Washington, DC.

Hamaker, E.L., Kuiper, R.M., & Grasman, R.P.P.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20,* 102-116.

Kenny, D.A. & Zautra, A. (2001). Trait-state models for longitudinal data. In L. Collins, & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 243-263), Washington, DC: American Psychological Association.

McArdle, J.J. & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. Collins, & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 137-175). Washington, DC: American Psychological Association.

Usami, S., Hayes, T., & McArdle, J.J. (2015). On the mathematical relationship between latent change score model and autoregressive cross-lagged factor approaches: Cautions for inferring causal relationship between variables. *Multivariate Behavioral Research, 50*, 676-687.

SATOSHI USAMI; usami@ct.u-tokyo.ac.jp