STRUCTURAL EQUATION MODELING
2025, VOL. 32, NO. 1
ISSN: 1070-5511 | E-ISSN: 1532-8007

OPEN ACCESS

A Note on the Occurrence of the Illusory Between-Person Component in the Random Intercept Cross-Lagged Panel Model

Alexander Robitzsch[a,b] and Oliver Lüdtke[a,b]

[a]IPN – Leibniz Institute for Science and Mathematics Education; [b]Centre for International Student Assessment (ZIB)

ABSTRACT
The random intercept cross-lagged panel model (RICLPM) decomposes longitudinal associations between two processes *X* and *Y* into stable between-person associations and temporal within-person changes. In a recent study, Bailey et al. demonstrated through a simulation study that the between-person variance components in the RICLPM can occur only due to the presence of time-varying covariate processes that are omitted from the analysis model. Therefore, the between-person component was termed illusory. In this article, necessary and sufficient conditions for the occurrence of such illusory between-person components in an RICLPM are analytically derived and systematically investigated in a simulation study.

KEYWORDS
Cross-lagged effects; illusory between-person component; panel data; random intercept cross-lagged panel model; stable factor

1. Introduction

In many areas of psychological research, the cross-lagged panel model (CLPM) is applied to estimate the effect of one variable X on another variable Y across time (i.e., cross-lagged effect). As an extension of the traditional CLPM, Hamaker et al. (2015) proposed the random intercept CLPM (RICLPM) that allows the decomposition of the longitudinal associations between two variables X and Y into stable between-person associations and temporal within-person changes. One particular feature of the RICLPM that many researchers find useful (Bailey et al., 2020; Lucas, 2023; Murayama & Gfrörer, 2022; Usami et al., 2019) is that the RICLPM provides within-person cross-lagged effects that are adjusted for the effects of stable between-person components (also labeled as stable trait factors).

Recently, Bailey et al. (2023) demonstrated through a simulation study that substantial between-person variance components can occur in a bivariate RICLPM only from the presence of time-varying covariate processes that are omitted from the analysis model. More specifically, their simulation results showed that the RICLPM provides significant estimates of stable trait variances, even if no stable trait factors were included in the data-generating model. Thus, the stable trait factors in the RICLPM also seem to absorb the effects of omitted time-varying covariates. Based on these results, Bailey et al. (2023, p. 2) concluded that stable trait variances in the RICLPM "can emerge from causal dynamics of time varying-processes, which are omitted from the analysis model, potentially leading to the estimation of traits that are, at least in part, illusory."

The possible presence of illusory traits in the RICLPM challenges the common interpretation that the RICLPM controls for (time-invariant) stable trait factors when estimating cross-lagged effects and needs further investigation. In the present article, we thoroughly study the conditions under which such illusory traits emerge in the RICLPM. The extent of the estimated between-person variance is analytically derived as a function of the (omitted) time-varying covariate process using some approximating assumptions for a longitudinal design with three measurement waves. Moreover, the analytically obtained equations for the estimated between-person variance are confirmed in a simulation study that also extends the results to different longitudinal designs. Therefore, while our study replicates the findings of Bailey et al. (2023) regarding the illusory trait on the one hand, it also states the conditions under which such a trait occurs on the other hand.

The rest of this article is structured as follows. Section 2 presents estimating equations of the model parameters for the univariate and the multivariate RICLPM for three time points. In Section 3, we derive the estimated variance in a univariate RICLPM if the data-generating model includes a time-varying target variable and a time-varying covariate process. Section 4 illustrates the analytical findings of a small simulation study. Finally, the paper concludes with a discussion in Section 5.

2. Random Intercept Cross-Lagged Panel Model

In this section, we review the RICLPM and present estimating equations for the model parameters. It is shown that the model parameters are just identified for a stationary process with three time points. In the following, the details for the univariate RICLPM are presented in Section 2.1. Afterward, the estimating equations are generalized to the case of the multivariate RICLPM in Section 2.2.

2.1. Univariate Random Intercept Cross-Lagged Panel Model

Assume that a target variable Y is longitudinally measured at three time points. The observation for person i at time t is denoted as Y_{it}, where $i = 1, \ldots, N$ denote persons and $t = 0, \ldots, T$ denote time points. The univariate RICLPM imposes a within-between decomposition for observations Y_{it}:

$$Y_{it} = Y_i^B + Y_{it}^W, \quad (1)$$

where the variances are denoted by $\mathrm{Var}(Y_i^B) = \tau_Y$ and $\mathrm{Var}(Y_{it}^W) = \sigma_Y$. Without loss of generality, we assume that the process has zero means (i.e., $\mathrm{E}(Y_{it}) = 0$ for all $t = 0, \ldots, T$). The time-invariant between-person part Y_i^B and the time-varying within-person part Y_{it}^W are assumed to be uncorrelated. The random variable Y_i^B is called the random intercept and also referred to as the between-person component (also called the stable trait), and τ_Y is frequently labeled as the between-person variance (variance of stable trait factor). Note that the random intercept Y_i^B does not vary across time. In contrast, the variable Y_{it}^W (i.e., within residuals) describes the variation of person i around its intercept Y_i^B and the corresponding variance σ_Y is referred to as the within-person variance. The univariate RICLPM[1] imposes an autoregressive process of order one, which is defined as

$$Y_{it}^W = \beta_Y Y_{i,t-1}^W + e_{Yit} \text{ for } t = 1, \ldots, T, \quad (2)$$

where β_Y is the autoregressive coefficient and e_{Yit} is a residual variable. The variance of the residual variable is given as $\mathrm{Var}(e_{Y,it}) = \sigma_Y(1 - \beta_Y^2)$. Figure 1 depicts the univariate RICLPM. By using the model Equations (1) and (2), the lagged covariance for a time lag $h \geq 0$ of the univariate process Y_{it} is given as

$$s_h = \mathrm{Cov}(Y_{it}, Y_{i,t-1}) = \tau_Y + \beta_Y^h \sigma_Y \text{ for } h = 0, \ldots, T. \quad (3)$$

We now derive the estimating equations for three time points. For a stationary process, three autocovariances s_0, s_1, and s_2 are sufficient statistics for the model parameters τ_Y, σ_Y, and β_Y. The equations for the three autocovariances defined in Equation (3) are nonlinear functions of the model parameters. A simple algebraic calculation provides (see Lüdtke & Robitzsch, 2022)

$$s_1 - s_0 = (\beta_Y - 1)\sigma_Y \quad \text{and}$$
$$s_2 - s_1 = \beta_Y(\beta_Y - 1)\sigma_Y = \beta Y(s_1 - s_0). \quad (4)$$

Hence, we get the estimating equations for the model parameters of the RICLPM from Equations (3) and (4) as

$$\beta_Y = \frac{s_2 - s_1}{s_1 - s_0}, \quad (5)$$

$$\sigma_Y = \frac{s_1 - s_0}{\beta_Y - 1} = \frac{(s_1 - s_0)^2}{s_2 - 2s_1 + s_0}, \quad (6)$$

and

$$\tau_Y = s_0 - \sigma_Y = s_0 - \frac{(s_1 - s_0)^2}{s_2 - 2s_1 + s_0}. \quad (7)$$

For more than three time points, the model is overidentified. However, the autocovariances s_h ($h = 0, \ldots, T$) are still sufficient statistics. An estimation method must be defined in order to express the parameters τ_Y, σ_Y, and β_Y as a function of autocovariances s_h. Estimating equations can be derived for, e.g., unweighted least squares estimation. However, the formulas typically provide less insight than in the just identified case of three time points. As the main motivation of this article is to demonstrate that a substantial between-person variance component can occur only due to (omitted) time-varying covariate processes, it is sufficient to analyze the phenomenon in the simplest case of three time points. In the following subsection, we discuss the multivariate RICLPM and its estimating equations.

2.2. Multivariate Random Intercept Cross-Lagged Panel Model

First, assume that there are two time-varying processes Y_{it} and X_{it}. The RICLPM for two processes (i.e., the bivariate RICLPM) also starts by decomposing the processes into a time-invariant between-person part and a time-varying within-person part (Hamaker et al., 2015).

$$Y_{it} = Y_i^B + Y_{it}^W$$
$$X_{it} = X_i^B + X_{it}^W \quad (8)$$

As in the univariate case, an autoregressive process of order one on the within residual variables Y_{it}^W and X_{it}^W is imposed

$$Y_{it}^W = \beta_{YY} Y_{i,t-1}^W + \beta_{YX} X_{i,t-1}^W + e_{Yit}$$
$$X_{it}^W = \beta_{XY} Y_{i,t-1}^W + \beta_{XX} X_{i,t-1}^W + e_{Xit}, \quad (9)$$

where β_{YY} and β_{XX} denote autoregressive coefficients, and β_{YX} and β_{XY} denote the cross-lagged coefficients. Figure 2 depicts the bivariate RICLPM.

The estimating equations for the model parameters can again be derived for a stationary process and three time points. To avoid burdensome notation, we consider a

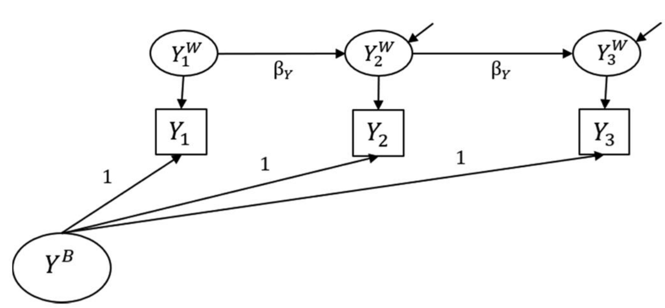

Figure 1. Path diagram of the univariate random intercept cross-lagged panel model for a process Y_{it}.

[1]Semantically, the label „univariate RICLPM" is a misnomer because the model does not include cross-lagged effects. However, we nevertheless use this label because it eases the generalization to the multivariate RICLPM and is also used in the literature (Andersen, 2022).

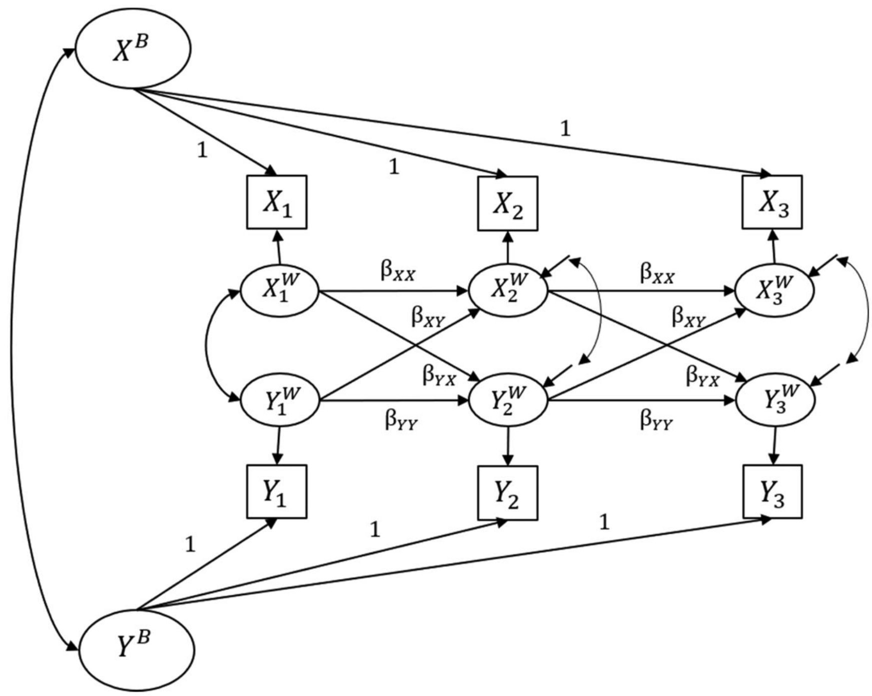

Figure 2. Path diagram of the bivariate random intercept cross-lagged panel model for processes Y_{it} and X_{it}.

multivariate RICLPM process $\mathbf{Y}_{it}$. The process is decomposed into uncorrelated between-person and within-person parts

$$\mathbf{Y}_{it} = \mathbf{Y}_i^B + \mathbf{Y}_{it}^W \tag{10}$$

with variance matrices $\mathrm{Var}(\mathbf{Y}_i^B) = \mathbf{T}_\mathbf{Y}$ and $\mathrm{Var}(\mathbf{Y}_{it}^W) = \mathbf{\Sigma}_\mathbf{Y}$. The autoregressive process of order one is defined for within residual variables $\mathbf{Y}_{it}^W$ and is given as (see Lüdtke & Robitzsch, 2022)

$$\mathbf{Y}_{it}^W = \mathbf{B}_\mathbf{Y}\mathbf{Y}_{i,t-1}^W + \mathbf{e}_{\mathbf{Y}it}, \tag{11}$$

where the coefficient matrix $\mathbf{B}_\mathbf{Y}$ contains the autoregressive and cross-lagged regression coefficients. Now, the autocovariance matrices $\mathbf{S}_h$ for a time lag h can be computed as

$$\mathbf{S}_h = \mathrm{Cov}(\mathbf{Y}_{it}, \mathbf{Y}_{i,t-1}) = \mathbf{T}_\mathbf{Y} + \mathbf{B}_\mathbf{Y}^h\mathbf{\Sigma}_\mathbf{Y} \tag{12}$$

As in the univariate case, the model parameters $\mathbf{T}_\mathbf{Y}$, $\mathbf{\Sigma}_\mathbf{Y}$, and $\mathbf{B}_\mathbf{Y}$ can be expressed as a function of the three observed autocovariance matrices $\mathbf{S}_0$, $\mathbf{S}_1$, and $\mathbf{S}_2$. By simple matrix computations, we obtain

$$\mathbf{S}_1 - \mathbf{S}_0 = (\mathbf{B}_\mathbf{Y} - \mathbf{I})\mathbf{\Sigma}_\mathbf{Y} \tag{13}$$

and

$$\mathbf{S}_2 - \mathbf{S}_1 = \mathbf{B}_\mathbf{Y}(\mathbf{B}_\mathbf{Y} - \mathbf{I})\mathbf{\Sigma}_\mathbf{Y} = \mathbf{B}_\mathbf{Y}(\mathbf{S}_1 - \mathbf{S}_0), \tag{14}$$

where $\mathbf{I}$ denotes the identity matrix. The unknown model parameters can be uniquely estimated from the lagged covariance matrices $\mathbf{S}_h$ by

$$\mathbf{B}_\mathbf{Y} = (\mathbf{S}_2 - \mathbf{S}_1)(\mathbf{S}_1 - \mathbf{S}_0)^{-1}, \tag{15}$$

$$\mathbf{\Sigma}_\mathbf{Y} = (\mathbf{B}_\mathbf{Y} - \mathbf{I})^{-1}(\mathbf{S}_1 - \mathbf{S}_0) = (\mathbf{S}_1 - \mathbf{S}_0)^2(\mathbf{S}_2 - 2\mathbf{S}_1 + \mathbf{S}_0)^{-1}, \tag{16}$$

and

$$\mathbf{T}_\mathbf{Y} = \mathbf{S}_0 - \mathbf{\Sigma}_\mathbf{Y} = \mathbf{S}_0 - (\mathbf{S}_1 - \mathbf{S}_0)^2(\mathbf{S}_2 - 2\mathbf{S}_1 + \mathbf{S}_0)^{-1}. \tag{17}$$

The estimating equations for the general case of more than three time points cannot be expressed in closed form. The extension to alternative longitudinal designs is further studied in the simulation.

3. Illusory Between-Person Component Captures Time-Varying Covariates

In this section, we investigate the conditions in which the estimated between-person variance in a univariate RICLPM captures the effects of an omitted time-varying covariate. Bailey et al. (2023) simulated a data-generating model (DGM) with a target process Y_{it} and a time-varying multivariate covariate process $\mathbf{Z}_{it}$ that are related to each other and follow a multivariate RICLPM. In order to enable deriving an insightful formula for the between-person variance, we first discuss the case of a univariate time-varying covariate Z_{it}. In what follows, the DGM is a bivariate RICLPM involving the two processes Y_{it} and Z_{it}. The analysis model of interest is the univariate RICLPM for the process Y_{it}. The bivariate RICLPM also decomposes Y_{it} and Z_{it} into a between-person part and a within-person part

$$Y_{it} = Y_i^B + Y_{it}^W \quad \text{and}$$
$$Z_{it} = Z_i^B + Z_{it}^W. \tag{18}$$

An autoregressive process of order one is imposed for the variables at the within-level

$$Y_{it}^W = \beta_{YY}Y_{i,t-1}^W + \beta_{YZ}Z_{i,t-1}^W + e_{Yit}$$
$$Z_{it}^W = \beta_{ZY}Y_{i,t-1}^W + \beta_{ZZ}Z_{i,t-1}^W + e_{Zit} \tag{19}$$

The DGM is graphically depicted in Figure 3.

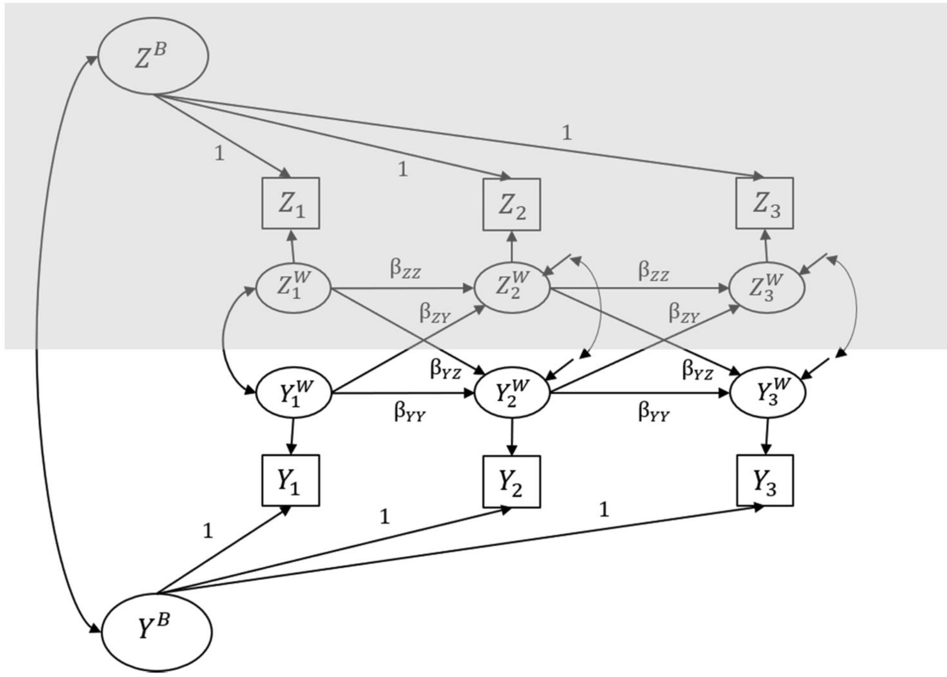

Figure 3. Path diagram of the bivariate random intercept cross-lagged panel model for processes Y_{it} and Z_{it}. The unobserved process Z_{it} is displayed with a gray-colored background.

In applications, the process Z_{it} might be considered a mediating process variable. Therefore, it is intentionally left from the analysis model, which is the univariate RICLPM for the process Y_{it} (depicted in Figure 1). The autoregressive model is given as

$$Y_{it}^W = \beta_Y Y_{i,t-1}^W + \tilde{e}_{yit} \tag{20}$$

The main interest is the size of the estimated variance $\tau_Y = \mathrm{Var}(Y_i^B)$ in the univariate RICLPM if the DGM involves the time-varying covariate process Z_{it}.

In Section 2.1, we derived the estimating equations for the model parameters (including the between-person variance τ_Y) in a univariate RICLPM. A simple method to derive the consequences of the DGM is to compute model-implied autocovariances for Y_{it} and to insert the quantities in the estimating equations. However, the resulting formulas are nonlinear and do not directly lead to insights. Therefore, we pursue an alternative approach for the derivation in this paper. The estimated between-person variance is derived under the assumption of "small" associations between the processes Y_{it} and Z_{it}. That is, we linearize the estimating equations for sufficiently small cross-lagged coefficients β_{YZ} and β_{ZY} such that the linear approximation is an acceptable approximation of the nonlinear estimating equation. Note that small values of β_{YZ} and β_{ZY} imply small changes in the observed covariances s_h $(h = 0, 1, 2)$.

More formally, denote by $\mathbf{s}$ the vector of lagged covariances of the process Y_{it} and by $\boldsymbol{\theta}$ the vector that includes the model parameters τ_Y, σ_Y, and β_Y. The estimating equations can be generally written as a solution of the multivariate nonlinear equation involving a multivariate function H such that (see Boos & Stefanski, 2013)

$$H(\mathbf{s}, \boldsymbol{\theta}) = 0 \tag{21}$$

Now, assume that there are lagged covariances $\mathbf{s}^*$ and a parameter $\boldsymbol{\theta}^*$ in the situation in which there is no relation of Y_{it} to the process Z_{it} (i.e., $\beta_{YZ} = 0$ and $\beta_{ZY} = 0$). We can linearize the estimation problem Equation (21) by defining $\mathbf{s} = \mathbf{s} + \Delta\mathbf{s}$ and $\boldsymbol{\theta} = \boldsymbol{\theta} + \Delta\boldsymbol{\theta}$ such that

$$\begin{aligned} 0 &= H(\mathbf{s}^* + \Delta\mathbf{s}, \boldsymbol{\theta}^* + \Delta\boldsymbol{\theta}) \\ &\cong H(\mathbf{s}^*, \boldsymbol{\theta}^*) + H_{\mathbf{s}}(\mathbf{s}^*, \boldsymbol{\theta}^*)\Delta\mathbf{s} + H_{\boldsymbol{\theta}}(\mathbf{s}^*, \boldsymbol{\theta}^*)\Delta\boldsymbol{\theta}, \end{aligned} \tag{22}$$

where $H_{\mathbf{s}}$ and $H_{\boldsymbol{\theta}}$ denote partial derivatives of H with respect to $\mathbf{s}$ and $\boldsymbol{\theta}$. Due to the definition $H(\mathbf{s}^*, \boldsymbol{\theta}^*) = 0$, the change in the estimated model parameter $\Delta\boldsymbol{\theta}$ can be determined as a function of the change in the autocovariance $\Delta\mathbf{s}$ as

$$\Delta\boldsymbol{\theta} \cong -H_{\boldsymbol{\theta}}(\mathbf{s}^*, \boldsymbol{\theta}^*)^{-1} H_{\mathbf{s}}(\mathbf{s}^*, \boldsymbol{\theta}^*)\Delta\mathbf{s}. \tag{23}$$

Equation (23) provides a formula for a parameter estimate by linearizing the nonlinear estimation problem.

We now apply this technique to estimate the between-person variance τ_Y for the process Y in a univariate RICLPM in the presence of Z_{it} in the DGM. The estimating equations for the univariate RICLPM defining the function H are presented in Section 2.1. By taking the partial derivatives of H with respect to $\mathbf{s} = (s_0, s_1, s_2)$ and $\boldsymbol{\theta} = (\tau_Y, \sigma_Y, \beta_Y)$, we get from Equation (3)

$$\begin{aligned} \Delta s_0 &= \Delta\tau_Y + \Delta\sigma_Y \\ \Delta s_1 &= \Delta\tau_Y + \beta_Y\Delta\sigma_Y + \sigma_Y\Delta\beta_Y \\ \Delta s_2 &= \Delta\tau_Y + \beta_Y^2\Delta\sigma_Y + 2\beta_Y\sigma_Y\Delta\beta_Y \end{aligned} \tag{24}$$

Because we can assume that the process Y_{it} has a constant variance of one, we have $\Delta s_0 = 0$. Therefore, the first equation in Equation (24) implies

$$\Delta\tau_Y = -\Delta\sigma_Y \tag{25}$$

Hence, Equation (24) can be simplified to

$$\begin{aligned}\Delta s_1 &= \Delta\tau_Y(1-\beta_Y) + \sigma_Y\Delta\beta_Y\\ \Delta s_2 &= \Delta\tau_Y(1-\beta_Y)(1+\beta_Y) + 2\beta_Y\sigma_Y\Delta\beta_Y\end{aligned} \tag{26}$$

Moreover, we obtain from Equation (26)

$$\Delta s_2 - 2\beta_Y\Delta s_1 = (1-\beta_Y)^2\ \Delta\tau_Y, \tag{27}$$

which results in

$$\Delta\tau_Y = \frac{\Delta s_2 - 2\beta_Y\Delta s_1}{(1-\beta_Y)^2}. \tag{28}$$

We can now compute the lagged covariances under the DGM of the processes Y_{it} and Z_{it} as

$$\begin{aligned}s_1 &= \tau_Y + \beta_{YY}\sigma_Y + \beta_{YZ}\sigma_{YZ}\\ s_2 &= \tau_Y + (\beta_{YY}^2 + \beta_{YZ}\beta_{ZY})\sigma_Y + \beta_{YZ}(\beta_{YY}+\beta_{ZZ})\sigma_{YZ}\end{aligned} \tag{29}$$

The changes in autocovariances can be written as

$$\begin{aligned}\Delta s_1 &= \beta_{YZ}\sigma_{YZ}\\ \Delta s_2 &= \beta_{YZ}\beta_{ZY}\sigma_Y + \beta_{YZ}(\beta_{YY}+\beta_{ZZ})\sigma_{YZ}\end{aligned} \tag{30}$$

Using Equations (28) and (30), the change in the estimated between-person variance can be expressed as (using $\beta_Y = \beta_{YY}$)

$$\Delta\tau_Y = \beta_{YZ}\frac{(\beta_{ZZ}-\beta_{YY})\sigma_{YZ} + \beta_{ZY}\sigma_Y}{(1-\beta_{YY})^2}. \tag{31}$$

Formula (31) for the change in the estimated between-person variance $\Delta\tau_Y$ provides conditions under which time-varying aspects of the covariate process Z_{it} generate between-person variance τ_Y for the process Y_{it} in the univariate RICLPM. As a necessary requirement, there must be a path from $Z_{i,t-1}$ to Y_{it} to ensure $\beta_{YZ} > 0$. Hence, the process Z_{it} influences the process Y_{it}. The denominator in (31) contains two terms that allow the change of the between variance $\Delta\tau_Y$ to be positive. First, the term $(\beta_{ZZ}-\beta_{YY})\sigma_{YZ}$ will be positive (a positive covariance σ_{YZ} is likely fulfilled in applications), if $\beta_{ZZ}-\beta_{YY}$ is larger than zero. That is, the (conditional) autoregressive stability of the process Z_{it} is larger than the counterpart for the process Y_{it}. Second, the term $\beta_{ZY}\sigma_Y$ is positive if the cross-lagged coefficient β_{ZY} is positive. That is, there is a path from $Y_{i,t-1}$ to Z_{it} that ensures $\beta_{ZY} > 0$. To summarize, the necessary condition $\beta_{YZ} > 0$ and one of the conditions $\beta_{ZZ} > \beta_{YY}$ or $\beta_{ZY} > 0$ must be fulfilled in order to estimate a positive between-person variance of Y_{it} that emerges from time-varying aspects of a covariate process Z_{it} and its relation to Y_{it}.

As for the between variance, changes in the autoregressive coefficient β_Y can also be derived as a function of the parameters of the DGM. Similar calculations for $\Delta\beta_Y$ result in

$$\Delta\beta_Y = \beta_{YZ}\frac{(1-\beta_{ZZ})\sigma_{YZ} - \beta_{ZY}\sigma_Y}{\sigma_Y(1-\beta_{YY})}. \tag{32}$$

3.1. Longer Time Lag $\Delta t = K$ Between Measurements

In the previous section, we derived the estimated between-person variance $\Delta\tau_Y$ in the univariate RICLPM if measurements at $t=0$, $t=1$, and $t=2$ are available at the time interval $[0,T]=[0,2]$. We now study how the estimated between variance changes if we consider three observations at the time interval $[0,T]=[0,2K]$ for an integer $K>1$. That is, a time lag of $\Delta t = K$ between measurements is assumed. The estimated between-person component variance $\Delta\tau_Y^{(K)}$ now depends on the time lag K, and formula (31) can be rewritten as

$$\Delta\tau_Y^{(K)} = \beta_{YZ}^{(K)}\frac{\left(\beta_{ZZ}^{(K)} - \beta_{YY}^{(K)}\right)\sigma_{YZ} + \beta_{ZY}^{(K)}\sigma_Y}{\left(1-\beta_{YY}^{(K)}\right)^2}, \tag{33}$$

where regression coefficients corresponding to time lag K are now involved in Equation (33). The matrix of coefficients of the autoregressive process for the time lag K is given by $\mathbf{B}^K$ if $\mathbf{B}$ denotes the coefficient matrix for a time lag of 1. By ignoring higher-order terms in β_{YZ} and β_{ZY}, we get

$$\begin{aligned}&\beta_{YY}^{(K)} \cong \beta_{YY}^K, \beta_{ZZ}^{(K)} \cong \beta_{ZZ}^K,\\ &\beta_{YZ}^{(K)} \cong \beta_{YZ}\sum\nolimits_{k=0}^{K-1}\beta_{YY}^k\beta_{ZZ}^{K-k-1},\end{aligned}$$

and

$$\beta_{ZY}^{(K)} \cong \beta_{ZY}\sum\nolimits_{k=0}^{K-1}\beta_{YY}^k\beta_{ZZ}^{K-k-1}. \tag{34}$$

We can now use

$$\beta_{ZZ}^{(K)} - \beta_{YY}^{(K)} = (\beta_{ZZ}-\beta_{YY})\sum\nolimits_{k=0}^{K-1}\beta_{YY}^k\beta_{ZZ}^{K-k-1} \tag{35}$$

and obtain

$$\Delta\tau_Y^{(K)} = \left[\frac{(1-\beta_{YY})\sum_{k=0}^{K-1}\beta_{YY}^k\beta_{ZZ}^{K-k-1}}{1-\beta_{YY}^K}\right]^2\Delta\tau_Y. \tag{36}$$

The estimated variance in Equation (36) can be bounded be

$$\left|\Delta\tau_Y^{(K)}\right| \le K^2\tilde{\beta}^{2(K-1)}|\Delta\tau_Y|, \tag{37}$$

where $\tilde{\beta} = \max(\beta_{YY}, \beta_{ZZ})$. Note that the factor in Equation (37) containing K converges to zero if the time lag K tends to infinity. Hence, the estimated between-person variance vanishes for a sufficiently long time interval $[0,T]=[0,2K]$.

3.2. Multivariate Covariate Process Z_{it}

Researchers Bailey et al. (2023) simulated a multivariate covariate process $\mathbf{Z}_{it}$ related to Y_{it} instead of a univariate process Z_{it}, which was treated in the previous section. Fortunately, it turns out that our findings for the univariate case do not substantially differ from the multivariate case. Assume a within-between decomposition for the processes Y_{it} and $\mathbf{Z}_{it}$ such that

$$\begin{aligned}Y_{it} &= Y_i^B + Y_{it}^W\\ \mathbf{Z}_{it} &= \mathbf{Z}_i^B + \mathbf{Z}_{it}^W.\end{aligned} \tag{38}$$

The DGM for the autoregressive process of order one operates on the random variables of the within-person part and is given by

$$Y_{it}^{W} = \beta_{YY} Y_{i,t-1}^{W} + \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top} \mathbf{Z}_{i,t-1}^{W} + e_{Yit}$$
$$\mathbf{Z}_{it}^{W} = \boldsymbol{\beta}_{\mathbf{Z}Y} Y_{i,t-1}^{W} + \mathbf{B}_{\mathbf{ZZ}} \mathbf{Z}_{i,t-1}^{W} + \mathbf{e}_{\mathbf{Z}it}, \quad (39)$$

where $\boldsymbol{\beta}_{Y\mathbf{Z}}$ and $\boldsymbol{\beta}_{\mathbf{Z}Y}$ are column vectors, and $\mathbf{B}_{\mathbf{ZZ}}$ is a regression coefficient matrix. Moreover, let $\boldsymbol{\sigma}_{Y\mathbf{Z}} = \text{Cov}(Y_{it}^{W}, \mathbf{Z}_{it}^{W})$ be the column vector of the within-covariances of Y_{it} and $\mathbf{Z}_{it}$. We can now compute the lagged covariances under the DGM defined in Equation (39)

$$s_1 = \tau_Y + \beta_{YY}\sigma_Y + \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\boldsymbol{\sigma}_{Y\mathbf{Z}}$$
$$s_2 = \tau_Y + (\beta_{YY}^2 + \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\boldsymbol{\beta}_{\mathbf{Z}Y})\sigma_Y + (\beta_{YY}\boldsymbol{\beta}_{Y\mathbf{Z}}^{\top} + \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\mathbf{B}_{\mathbf{ZZ}})\boldsymbol{\sigma}_{Y\mathbf{Z}} \quad (40)$$

The Taylor approximation relies on small associations of Y_{it} and $\mathbf{Z}_{it}$.(i.e., $\boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}$ and $\boldsymbol{\beta}_{\mathbf{Z}Y}$ are sufficiently"small"). The changes in model-implied autocovariances for Y_{it} are computed as

$$\Delta s_1 = \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\boldsymbol{\sigma}_{Y\mathbf{Z}} \quad \text{and}$$
$$\Delta s_2 = \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\boldsymbol{\beta}_{\mathbf{Z}Y}\sigma_Y + \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}(\beta_{YY}\mathbf{I} + \mathbf{B}_{\mathbf{ZZ}})\boldsymbol{\sigma}_{Y\mathbf{Z}}, \quad (41)$$

where $\mathbf{I}$ is the identity matrix of the dimension of the process $\mathbf{Z}_{it}$. Hence, the change in the estimated between-person variance can be derived as

$$\Delta\tau_Y = \boldsymbol{\beta}_{Y\mathbf{Z}}^{\top}\frac{(\mathbf{B}_{\mathbf{ZZ}} - \beta_{YY}\mathbf{I})\boldsymbol{\sigma}_{Y\mathbf{Z}} + \boldsymbol{\beta}_{\mathbf{Z}Y}\sigma_Y}{(1-\beta_{YY})^2}. \quad (42)$$

Note that the multivariate case (42) is a direct generalization of the univariate case (31). To put it plainly, many covariate processes $\mathbf{Z}_{it}$ can have the same impact on the estimated between-person variance in the univariate RICLPM for Y_{it} like a single covariate process Z_{it}. One observes that at least one component of the multivariate process $\mathbf{Z}_{it}$ must impact Y_{it} in order to ensure $\boldsymbol{\beta}_{Y\mathbf{Z}}^{\top} = 0$.

3.3. Multivariate RICLPM $\mathbf{Y}_{it}$ and Multivariate Covariate Process $\mathbf{Z}_{it}$.

In this subsection, we generalize the findings of Sections 3.1. and 3.2. to the situation of multivariate processes $\mathbf{Y}_{it}$ and $\mathbf{Z}_{it}$. The following analysis is concerned with the estimation of the between-person variance matrix $\mathbf{T}_{\mathbf{Y}} = \text{Var}(\mathbf{Y}_i^B)$ of the between-person part $\mathbf{Y}_i^B$ of the process $\mathbf{Y}_{it}$. It is demonstrated that an illusory between-person variance $\mathbf{T}_{\mathbf{Y}}$ can occur due to time-varying aspects of a covariate process $\mathbf{Z}_{it}$ that is associated with $\mathbf{Y}_{it}$. The derivation is formally equivalent to the simple case of univariate processes Y_{it} and Z_{it}, but involves slightly more clumsy matrix notation.

First, we study the estimating equations of the multivariate RICLPM for the process $\mathbf{Y}_{it}$. Let $\mathbf{B}_{\mathbf{Y}}$ denote the matrix of autoregression coefficients of the within-person process $\mathbf{Y}_{it}^{W}$. We linearize the estimating equations (see Section 2.2) around small changes in autocovariances $\mathbf{S}_h$ (for $h=0,1,2$) to investigate changes in the between-person variance matrix $\boldsymbol{\Sigma}_{\mathbf{Y}}$ and the regression coefficients $\mathbf{B}_{\mathbf{Y}}$. As in the univariate case, we fix the cross-sectional covariance such that $\Delta\mathbf{T}_{\mathbf{Y}} = -\Delta\boldsymbol{\Sigma}_{\mathbf{Y}}$, where $\boldsymbol{\Sigma}_{\mathbf{Y}}$ denotes the within-person variance matrix. Under this assumption, the estimating equations for changes in autocovariances can be derived using matrix calculus (Magnus & Neudecker, 2019)

$$\Delta\mathbf{S}_1 = (\mathbf{I} - \mathbf{B}_{\mathbf{Y}})(\Delta\mathbf{T}_{\mathbf{Y}}) + (\Delta\mathbf{B}_{\mathbf{Y}})\boldsymbol{\Sigma}_{\mathbf{Y}}$$
$$\Delta\mathbf{S}_2 = \left(\mathbf{I} - \mathbf{B}_{\mathbf{Y}}^2\right)(\Delta\mathbf{T}_{\mathbf{Y}}) + (\Delta\mathbf{B}_{\mathbf{Y}})\mathbf{B}_{\mathbf{Y}}\boldsymbol{\Sigma}_{\mathbf{Y}} + \mathbf{B}_{\mathbf{Y}}(\Delta\mathbf{B}_{\mathbf{Y}})\boldsymbol{\Sigma}_{\mathbf{Y}} \quad (43)$$

Now, assume that a multivariate RICLPM holds as the DGM for the joint process $(\mathbf{Y}_{it}, \mathbf{Z}_{it})$. The regression coefficients of the within-person processes are given by

$$\mathbf{Y}_{it}^{W} = \mathbf{B}_{\mathbf{YY}}\mathbf{Y}_{i,t-1}^{W} + \mathbf{B}_{\mathbf{YZ}}\mathbf{Z}_{i,t-1}^{W} + \mathbf{e}_{\mathbf{Y}it}$$
$$\mathbf{Z}_{it}^{W} = \mathbf{B}_{\mathbf{ZY}}\mathbf{Y}_{i,t-1}^{W} + \mathbf{B}_{\mathbf{ZZ}}\mathbf{Z}_{i,t-1}^{W} + \mathbf{e}_{\mathbf{Z}it}. \quad (44)$$

Furthermore, let $\boldsymbol{\Sigma}_{\mathbf{YY}}$ be the within-person variance matrix of $\mathbf{Y}_{it}$, and $\boldsymbol{\Sigma}_{\mathbf{ZY}}$ is the within-person covariance matrix of $\mathbf{Z}_{it}$ and $\mathbf{Y}_{it}$. For regression coefficient matrices $\mathbf{B}_{\mathbf{YZ}}$ and $\mathbf{B}_{\mathbf{ZY}}$ with entries that do not deviate too large from zero, changes in the autocovariance matrices for the process $\mathbf{Y}_{it}$ are calculated as

$$\Delta\mathbf{S}_1 = \mathbf{B}_{\mathbf{YZ}}\boldsymbol{\Sigma}_{\mathbf{ZY}}$$
$$\Delta\mathbf{S}_2 = \mathbf{B}_{\mathbf{YZ}}\mathbf{B}_{\mathbf{ZY}}\boldsymbol{\Sigma}_{\mathbf{YY}} + (\mathbf{B}_{\mathbf{YY}}\mathbf{B}_{\mathbf{YZ}} + \mathbf{B}_{\mathbf{YZ}}\mathbf{B}_{\mathbf{ZZ}})\boldsymbol{\Sigma}_{\mathbf{ZY}} \quad (45)$$

By inserting Equation (45) into (43) and solving for $(\Delta\mathbf{T}_{\mathbf{Y}})$, we obtain

$$(\Delta\mathbf{T}_{\mathbf{Y}}) = (\mathbf{I} - \mathbf{B}_{\mathbf{Y}})^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}(\mathbf{I} - \mathbf{B}_{\mathbf{Y}})^{-1}\boldsymbol{\Sigma}_{\mathbf{YY}}, \quad (46)$$

where the matrix $\boldsymbol{\mathcal{B}}$←governs the size of the estimated between-person variance matrix if the target process $\mathbf{Y}_{it}$ is related to the covariate process $\mathbf{Z}_{it}$. The matrix $\boldsymbol{\mathcal{B}}$ is given as

$$\boldsymbol{\mathcal{B}} = \mathbf{B}_{\mathbf{YZ}}\left[\mathbf{B}_{\mathbf{ZZ}}\boldsymbol{\Sigma}_{\mathbf{ZY}} - \boldsymbol{\Sigma}_{\mathbf{ZY}}\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}\mathbf{B}_{\mathbf{YY}}\boldsymbol{\Sigma}_{\mathbf{YY}} + \mathbf{B}_{\mathbf{ZY}}\boldsymbol{\Sigma}_{\mathbf{YY}}\right]. \quad (47)$$

Importantly, the matrix $\boldsymbol{\mathcal{B}}$ has the same structure as the denominator in the formula for $\Delta\tau_Y$ in the univariate (see Equation (31)). A necessary condition for the occurrence of a between-person variance $\mathbf{T}_{\mathbf{Y}}$ are nonzero (and typically positive) coefficients in $\mathbf{B}_{\mathbf{YZ}}$. The sufficient conditions in the univariate case were simple. However, they can be described in the multivariate case that either the term $\mathbf{B}_{\mathbf{ZZ}}\boldsymbol{\Sigma}_{\mathbf{ZY}} - \boldsymbol{\Sigma}_{\mathbf{ZY}}\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}\mathbf{B}_{\mathbf{YY}}\boldsymbol{\Sigma}_{\mathbf{YY}}$ or the term $\mathbf{B}_{\mathbf{ZY}}\boldsymbol{\Sigma}_{\mathbf{YY}}$ transport positive associations with the process $\mathbf{Y}_{it}$ in order to obtain positive between-person variances in the matrix $\mathbf{T}_{\mathbf{Y}}$. It should be emphasized that the occurrence of a between variance $\mathbf{T}_{\mathbf{Y}}$ (i.e., the presence of an illusory trait) implies changes in regression coefficients of the autoregressive paths (i.e., in the matrix $\mathbf{B}_{\mathbf{Y}}$). Hence, cross-lagged coefficients will consequently change if illusory traits occur in the RICLPM.

4. Simulation Study

In this simulation study, we demonstrate the adequacy of the formula for the estimated between-person variance (i.e., Equation (31)) for the case of a univariate covariate process Z_{it}. Overall, the simulation shows that a between-person variance for Y_{it} is estimated that emerges solely from the association of Y_{it} with the time-varying process Z_{it}. We also extend the findings to longitudinal designs with more measurement points.

4.1. Method

This simulation only considers estimated parameters at the population level. Sampling of persons does not provide further insights because our arguments are conceptual and pertain to the model choice that can be safely addressed for population-level data. As in Section 3.1., the DGM involves the two processes Y_{it} and Z_{it} that follow a bivariate RICLPM, but the analysis model is a univariate RICLPM for the process Y_{it}. Assume that the processes Y_{it} and Z_{it} are standardized (i.e., they have zero means and standard deviations of one) and stationary. They are decomposed into a between-person part Y_i^B (and Z_i^B) and a within-person part Y_{it}^W (and Z_{it}^W). Denote by $\mathbf{B}_1$ the matrix of coefficients of the autoregressive process at the within-level corresponding to the time lag $\Delta t = 1$:

$$\mathbf{B}_1 = \begin{pmatrix} \beta_{YY} & \beta_{YZ} \\ \beta_{ZY} & \beta_{ZZ} \end{pmatrix}. \tag{48}$$

This coefficient matrix can be converted to other time lags Δt by means of the matrix exponential function (Kuiper & Ryan, 2018)

$$\mathbf{B}_{\Delta t} = \exp(\mathbf{B}_0^* \Delta t), \tag{49}$$

where exp denotes the matrix exponential function, and $\mathbf{B}_0^*$ is the drift matrix of a corresponding continuous time model. Note that there is a bijective correspondence of $\mathbf{B}_{\Delta t}$ and $\mathbf{B}_0^*$, meaning that the continuous time model and the discrete cross-lagged panel model are statistically equivalent. In this simulation study, the drift matrix $\mathbf{B}_0^*$ is obtained by solving Equation (49) for a given regression coefficient $\mathbf{B}_1$. Afterward, the matrix $\mathbf{B}_{\Delta t}$ can be computed by inserting $\mathbf{B}_0^*$ and a specified Δt (such as $\Delta t = 0.5$) into Equation (49).

The entire simulation computes a model-implied covariance matrix at the population level. Hence, parameter estimates are investigated at the population level. Across all simulation conditions, we fixed the autoregressive coefficients β_{YY} to 0.4, the correlation (at the between-person level) of Y_i^B and Z_i^B at 0.3, and the correlation (at the within-person level) of Y_{it}^W and Z_{it}^W at 0.5.

Other parameters were varied across the conditions of the simulation. First, measurements were obtained from the time interval $[0, T]$, where T was chosen as 2, 4, 6, or 8. Second, the time lag Δt was varied across conditions. For a fixed time interval length T, the number of measurements was given as $\frac{T}{\Delta t} + 1$. For all lengths T of the time interval $[0, T]$, we chose $\Delta t = 0.5$ or $\Delta t = 1$. Moreover, we chose Δt as 2, 3, and 4 in the conditions $T=$ 4, 6, and 8, respectively, such that there were conditions with three measurements for all values of T included. Third, the between variance of Z_i^B (i.e., τ_Z) was either 0 or 0.3. Fourth, the between variance of Y_i^B (i.e., τ_Y) was either 0 or 0.2. Fifth, the regression coefficient β_{YZ} was chosen as 0 or 0.3. According to the between-person variance formula (31), the choice $\beta_{YZ} = 0$ would result in a between-person variance of 0, while the condition $\beta_{YZ} = 0.3$ can allow a positively estimated between variance if one of two additional conditions is fulfilled. Sixth, the autoregressive coefficient β_{ZZ} was chosen as 0.4 or 0.7. The first condition implied $\beta_{ZZ} = \beta_{YY}$, while the second condition ensured $\beta_{ZZ} - \beta_{YY} > 0$ as part of a sufficient condition to a positively estimated between-person variance (see Equation (31)). Seventh, the regression coefficient β_{ZY} was chosen as either 0 or 0.2. According to Equation (31), the choice $\beta_{ZY} = 0.2$ will result in a positive between-person variance estimate in the case $\beta_{YZ} = 0.3$.

The univariate RICLPM was fitted in the R (R Core Team, 2023) package lavaan (Rosseel, 2012). The model was estimated with maximum likelihood and variance constraints on the between-person variance and the residual variance (i.e., larger than 0.0001) were imposed. Moreover, the stationarity conditions were implemented by specifying autoregression coefficients and residual variances as time-invariant. We also computed the model fit statistics (Jöreskog et al., 2016), root mean square error of approximation (RMSEA), and standardized root mean squared residual (SRMR) to assess the extent of model misspecification in the simulation conditions. Replication material for this simulation study can be found at https://osf.io/5qxk7/ (accessed on 1 July 2024).

4.2. Results

Table 1 displays the estimated between-person variance τ_Y for all simulation conditions. As predicted from the analytically derived formula (31), the regression coefficient β_{YZ} had to be larger than zero in order to estimate a positive between-person variance. In addition, the two additional sufficient conditions for a positive variance (i.e., either $\beta_{ZZ} > \beta_{YY}$ or $\beta_{ZY} > 0$) were also confirmed in the simulation. Notably, there were many conditions in which a large positive between-person variance was estimated, although no stable between variance τ_Y was simulated in the DGM. This confirms the simulation results of Bailey et al. (2023) that the RICLPM can falsely detect stable trait factors in the presence of omitted time-covariate processes. Overall, the estimated between variance was slightly larger for the time lag $\Delta t = 0.5$ than for $\Delta t = 1$. Moreover, in line with the predictions in Section 3.2, the estimated between variance decreased with a longer time interval $[0, T]$ in the conditions with three measurements. This finding was also observed for more than three time points with time lags Δt of 0.5 and 1. However, the estimated between-person variance decreased much more slowly with increasing time intervals $[0, T]$, but a fixed time lag Δt. Hence, the chance of observing an illusory between-person component (i.e., a positive between-person variance) is increased if more measurements are made. Importantly, the magnitude of the estimated between-person variance for a fixed time interval $[0, T]$ is also a function of the time lag Δt. The more measurements are made (i.e., the smaller the time lag Δt), the larger is the estimated between-person variance. The finding of a spuriously inflated between-person variance was also observed for a simulated between variance $\tau_Y = 0.2$. In these conditions, the estimated between-person variance was frequently larger than 0.2.

We also assessed model fit statistics in the conditions with $\beta_{YZ} = 0.3$ and a positively "biased" between-person

Table 1. Estimated variance of between-person variance τ_Y in a univariate RICLPM.

					$T=2$		$T=4$			$T=6$				$T=8$			
					Δt		Δt			Δt				Δt			
τ_Y	β_{YZ}	τ_Z	β_{ZY}	β_{ZZ}	0.5	1	0.5	1	2	0.5	1	2	3	0.5	1	2	4
0	0	0	0	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
			0.2	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		0.3	0	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
			0.2	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.3	0	0	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.11	0.10	0.09	0.08	0.06	0.07	0.06	0.05	0.03	0.06	0.05	0.04	0.02
			0.2	0.4	0.21	0.16	0.14	0.11	0.05	0.11	0.08	0.04	0.02	0.09	0.06	0.03	0.01
				0.7	0.31	0.29	0.28	0.26	0.21	0.25	0.23	0.19	0.14	0.22	0.20	0.16	0.09
		0.3	0	0.4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
				0.7	0.10	0.09	0.08	0.07	0.05	0.06	0.06	0.04	0.03	0.05	0.05	0.03	0.01
			0.2	0.4	0.21	0.16	0.15	0.11	0.06	0.12	0.09	0.04	0.02	0.09	0.07	0.03	0.01
				0.7	0.29	0.27	0.25	0.24	0.19	0.23	0.21	0.17	0.13	0.20	0.19	0.15	0.09
0.2	0	0	0	0.4	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
				0.7	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
			0.2	0.4	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
				0.7	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
		0.3	0	0.4	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
				0.7	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
			0.2	0.4	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
				0.7	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
	0.3	0	0	0.4	0.00	0.06	0.09	0.12	0.17	0.12	0.15	0.18	0.19	0.14	0.16	0.18	0.20
				0.7	0.30	0.28	0.28	0.27	0.25	0.26	0.25	0.24	0.22	0.25	0.25	0.23	0.21
			0.2	0.4	0.36	0.32	0.31	0.28	0.24	0.28	0.26	0.23	0.21	0.26	0.25	0.22	0.20
				0.7	0.46	0.44	0.43	0.42	0.38	0.41	0.39	0.36	0.32	0.39	0.37	0.34	0.28
		0.3	0	0.4	0.08	0.12	0.13	0.15	0.18	0.15	0.17	0.19	0.20	0.16	0.18	0.19	0.20
				0.7	0.29	0.28	0.27	0.26	0.24	0.26	0.25	0.24	0.22	0.25	0.24	0.23	0.21
			0.2	0.4	0.37	0.33	0.32	0.29	0.24	0.29	0.27	0.23	0.21	0.27	0.25	0.23	0.21
				0.7	0.44	0.42	0.42	0.40	0.36	0.39	0.38	0.35	0.31	0.37	0.36	0.33	0.27

Note. Conditions with three measurements are printed with a gray background.

variance. The maximum SRMR in these conditions was 0.034 ($M = 0.008$, $SD = 0.008$), and the maximum RMSEA was 0.027 ($M = 0.008$, $SD = 0.007$). Hence, these measures of absolute model fit were not indicative of an estimated artificial positive between-person variance that was a consequence of an omitted time-varying covariate process Z_{it}.

5. Discussion

This article investigates the conditions in which illusory between-person component variance estimates were obtained in a univariate RICLPM. Importantly, the estimated between-person variance was not due to stable differences between persons but was entirely caused by associations of a target variable Y_{it} with a time-varying process Z_{it}. The findings underscore the notion that a blind application of the RICLPM is not necessarily a safeguard to control stable factors within persons but can artificially capture parts of time-varying processes that should not be controlled because these could be potential mediators. Importantly, the occurrence of an illusory between variance will typically imply changes in estimated cross-lagged coefficients. Hence, researchers must consider the consequences of specifying a random intercept in cross-lagged analyses. Moreover, it has also been shown that model fit statistics were ineffective in detecting illusory trait variances.

For a fixed time interval, there is always a danger that the between-person component also includes time-varying elements of other process variables. The positive bias in the estimated between-person variance only vanishes if the length T of the time interval $[0, T]$ increases. However, increasing the number of measurements for a fixed time interval $[0, T]$ makes the situation even worse. We do not intend to argue that the RICLPM should never be used for studying cross-lagged relationships between process variables X_{it} and Y_{it}. However, in our opinion, there is no justification that researchers should automatically prefer the RICLPM over CLPM because the former model would provide less biased estimates of cross-lagged effects (see, e.g., Hamaker et al., 2015; Hamaker, 2023; Lucas, 2022, 2023; Schimmack, 2020). Hamaker (2023) pointed out that "between-person components merely capture stable factors with respect to the time span of a study, which implies that they are critically dependent on the study design that is used." This property has also been emphasized in the econometric literature (Mundlak, 1978; Millimet & Bellemare, 2023). Our article also underscores the fact that the extracted between-person components depend on the time interval and the time lag of measurements. It was highlighted that between-person component variance can occur due to the presence of a time-varying covariate process that does not contain stable components. Hence, as illustrated by Bailey et al. (2023), instable aspects of an (intentionally) omitted covariate process are captured in the between-person component variance. Therefore, we are unconvinced by general recommendations such that the RICLPM can "…account for stable factors—such as traits, trends, random intercepts, and unit effects—that represent lasting characteristics of individuals, thereby allowing for the investigation of within-person cross-lagged effects that are not contaminated by these stable between-person differences." (Hamaker, 2023; see also Schuurman, 2023). We believe that the term "stable factors" can only be metaphorically interpreted and does not refer to stable characteristics of persons. Hence, we think it is unwise to tell applied researchers and practitioners: "Even if the stable, between-person differences are not of interest, these nevertheless need to be separated from the within-person fluctuations. The stable mean differences between persons are often referred to as trait differences, but may also be conceptualized as resulting from stable confounders, such as stable genetic and environmental factors that affect our measurements." (Mulder & Hamaker, 2023). It would be more neutral to say that between-person components could potentially control stable confounders. However, as shown in our derivations and the simulation study, there is always a risk of overcontrolling time-varying unstable aspects of the RICLPM. Consequently, the call to abandon cross-lagged panel models without specified between-person components (see Schimmack, 2020; Lucas, 2022; but see Asendorpf, 2021, Lüdtke & Robitzsch, 2021, or Orth et al., 2021) is scientifically unjustified. Our cautionary remark regarding between-person factors also applies to extensions of the RICLPM model (e.g., Falkenström et al., 2023; Hori & Miyazaki, 2023; Lüdtke et al., 2023; Sorjonen et al., 2023).

We would also like to point out that the occurrence of an illusionary between-person component would also appear if continuous time models (CTSEMs; Hecht & Zitzmann, 2020; Lohmann et al., 2024; Oud, 2002; Voelkle et al., 2012) with a random intercept instead of discrete-time panel models, such as the RICLPM, were applied. For fixed time points across persons and an autoregressive process of order one, CTSEMs are statistically equivalent to discrete time models. Hence, the conceptual issues of what is captured by between-person components also occur in CTSEMs.

An anonymous reviewer wondered whether tests, sensitivity analyses or model comparisons could help to examine whether the random intercept variable is a true trait or just a reflection of omitted covariates. We do not think that model fit should be assessed or sensitivity analyses prove helpful because researchers have to theoretically define the covariates that should be controlled for if cross-lagged paths should be causally interpreted (Lüdtke & Robitzsch, 2023). Hence, even if a random intercept variance were statistically differing from zero, there is no general justification for specifying the RICLPM instead of CLPM if researchers cannot rule out the possibility that there are covariates implicitly controlled in the random intercept (i.e., the illusory trait) that should not be controlled for (but see Hamaker et al., 2015, for a different opinion). In this sense, referring to the cause of an illusory trait semantically as an omitted covariate is dangerous. While this is algebraically true by using our specified data-generating model in the analytical derivation and the simulation study, in causal inference questions, the target parameter of interest is defined as a functional of distribution of variables in the dataset, and researchers have to specify which variables should be controlled for (i.e., are confounders) and which variables are part of the causal effect of interest (i.e., are mediators or post-treatment variables). We believe that researchers are always more transparent and confident about what is controlled for if they rely on observed confounding variables in ordinary CLPM (possibly with autoregressive paths of higher order) instead of relying on inference based on unobserved confounding variables such as the latent random intercept variables in the RICLPM (Lüdtke & Robitzsch, 2022; see also Mulder et al., 2024).

References

Andersen, H. K. (2022). Equivalent approaches to dealing with unobserved heterogeneity in cross-lagged panel models? Investigating the benefits and drawbacks of the latent curve model with structured residuals and the random intercept cross-lagged panel model. *Psychological Methods*, *27*, 730–751. https://doi.org/10.1037/met0000285

Asendorpf, J. B. (2021). Modeling developmental processes. *In* J. R. Rauthmann (Ed.), *Handbook of personality dynamics and processes* (pp. 815–835). Elsevier. https://doi.org/10.1016/B978-0-12-813995-0.00031-5

Bailey, D. H., Hübner, N., Zitzmann, S., Hecht, M., & Murayama, K. (2023). Illusory traits: Wrong but sometimes useful. *PsyArXiv*. https://doi.org/10.31234/osf.io/rz8h4

Bailey, D. H., Oh, Y., Farkas, G., Morgan, P., & Hillemeier, M. (2020). Reciprocal effects of reading and mathematics? Beyond the cross-lagged panel model. *Developmental Psychology*, *56*, 912–921. https://doi.org/10.1037/dev0000902

Boos, D. D., & Stefanski, L. A. (2013). *Essential statistical inference: Theory and methods*. Springer. https://doi.org/10.1007/978-1-4614-4818-1

Falkenström, F., Solomonov, N., & Rubel, J. (2023). To detrend, or not to detrend, that is the question? The effects of detrending on cross-lagged effects in panel models. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000632

Hamaker, E. L. (2023). The within-between dispute in cross-lagged panel research and how to move forward. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000600

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. https://doi.org/10.1037/a0038889

Hecht, M., & Zitzmann, S. (2020). A computationally more efficient Bayesian approach for estimating continuous-time models. *Structural Equation Modeling*, *27*, 829–840. https://doi.org/10.1080/10705511.2020.1719107

Hori, K., & Miyazaki, Y. (2023). Latent curve detrending for disaggregating between-person effect and within-person effect. *Structural Equation Modeling*, *30*, 192–213. https://doi.org/10.1080/10705511.2022.2069113

Jöreskog, K. G., Olsson, U. H., & Wallentin, F. Y. (2016). *Multivariate analysis with LISREL*. Springer. https://doi.org/10.1007/978-3-319-33153-9

Kuiper, R. M., & Ryan, O. (2018). Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval. *Structural Equation Modeling*, *25*, 809–823. https://doi.org/10.1080/10705511.2018.1431046

Lohmann, J. F., Zitzmann, S., & Hecht, M. (2024). Studying between-subject differences in trends and dynamics: Introducing the random coefficients continuous-time latent curve model with structured residuals. *Structural Equation Modeling*, *31*, 151–164. https://doi.org/10.1080/10705511.2023.2192889

Lucas, R. E. (2022). It's time to abandon the cross-lagged panel model. *PsyArXiv*. https://doi.org/10.31234/osf.io/pkec7

Lucas, R. E. (2023). Why the cross-lagged panel model is almost never the right choice. *Advances in Methods and Practices in Psychological Science*, *6*, 1–22. https://doi.org/10.1177/25152459231158378

Lüdtke, O., & Robitzsch, A. (2021). A critique of the random intercept cross-lagged panel model. *PsyArXiv*. https://doi.org/10.31234/osf.io/6f85c

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling*, *29*, 888–907. https://doi.org/10.1080/10705511.2022.2065278

Lüdtke, O., & Robitzsch, A. (2023). ANCOVA versus change score for the analysis of two-wave data. *The Journal of Experimental Education*. Advance online publication. https://doi.org/10.1080/00220973.2023.2246187

Lüdtke, O., Robitzsch, A., & Ulitzsch, E. (2023). A Bayesian approach to estimating reciprocal effects with the bivariate STARTS model. *Multivariate Behavioral Research*, *58*, 560–579. https://doi.org/10.1080/00273171.2022.2039585

Magnus, J. R., & Neudecker, H. (2019). *Matrix differential calculus with applications in statistics and econometrics*. Wiley. https://doi.org/10.1002/9781119541219

Millimet, D., & Bellemare, M. F. (2023). Fixed effects and causal inference. *SSRN Electronic Journal*, 16202. https://doi.org/10.2139/ssrn.4467963

Mulder, J. D., & Hamaker, E. L. (2023). *The RI-CLPM & extensions. Frequently asked questions*. accessed on 23 August 2023. https://jeroendmulder.github.io/RI-CLPM/faq.html

Mulder, J. D., Usami, S., & Hamaker, E. L. (2024). Joint effects in cross-lagged panel research using structural nested mean models. *Structural Equation Modeling*. Advance online publication. https://doi.org/10.1080/10705511.2024.2355579

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*, 69–85. https://doi.org/10.2307/1913646

Murayama, K., & Gfrörer, T. (2022). Thinking clearly about time-invariant confounders in cross-lagged panel models: A guide for choosing a statistical model from a causal inference perspective. *PsyArXiv*. https://doi.org/10.31234/osf.io/bt9xr

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*, 1013–1034. https://doi.org/10.1037/pspp0000358

Oud, J. H. (2002). Continuous time modeling of the cross-lagged panel design. *Kwantitatieve Methoden*, *69*, 1–26.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Schimmack, U. (2020). *Why most cross-lagged models are false*. Retrieved August 23, 2023, from https://replicationindex.com/2020/08/22/cross-lagged/

Schuurman, N. K. (2023). A "within/between problem" primer: About (not) separating within-person variance and between-person variance in psychology. *PsyArXiv*. https://doi.org/10.31234/osf.io/7zgkx

Sorjonen, K., Nilsonne, G., Melin, B., & Ingre, M. (2023). Uncertain inference in random intercept cross-lagged panel models: An example involving need for cognition and anxiety and depression symptoms. *Personality and Individual Differences*, *201*, 111925. https://doi.org/10.1016/j.paid.2022.111925

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*, 637–657. https://doi.org/10.1037/met0000210

Voelkle, M. C., Oud, J. H., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*, 176–192. https://doi.org/10.1037/a0027543