Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The case of contemporaneous and reciprocal effects

Bengt Muthén and Tihomir Asparouhov *

December 1, 2023

*We thank Ellen Hamaker, Michael Zyphur, and David Kenny for comments on an earlier version. Noah Hastings provided support with tables and figures.

Abstract

This paper considers identification, estimation, and model fit issues for models with contemporaneous and reciprocal effects. It explores how well the models work in practice using Monte Carlo studies as well as real-data examples. Furthermore, by using models that allow contemporaneous and reciprocal effects, the paper raises a fundamental question about current practice for cross-lagged panel modeling using models such as CLPM or RI-CLPM: Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The paper concludes that the answer is no, a finding that has important ramifications for current practice. It is suggested that analysts should use additional models to probe the temporalities of the CLPM, RI-CLPM effects to see if these could be considered contemporaneous rather than lagged.

Keywords: panel data; equivalent models; nonrecursive models; lag0; random intercept; RI-CLPM; depression and self-esteem

1 Introduction

Panel data modeling with cross-lagged effects between two or more variables is a popular analysis technique especially in psychology. According to the overview article by Orth et al. (2021): "Cross-lagged regression models are by far the most commonly used method to test the prospective effect of one construct on another". Common approaches are the cross-lagged panel model (CLPM) and the random intercept cross-lagged panel model (RI-CLPM), but several other model variants are available (see, e.g., Asparouhov & Muthén, 2022; Hamaker et al., 2015; Hamaker, 2023; Orth et al., 2021; Zyphur et al., 2020). This paper goes beyond such traditional analysis techniques and considers panel data modeling with contemporaneous and reciprocal effects. In this paper, the term reciprocal effects refers to bi-directional contemporaneous (lag0) effects as opposed to the convention of calling cross-lagged effects (lag1, lag2, etc.) reciprocal. A key question is if models with reciprocal effects are identified. The answer is yes but most cross-lagged panel data analysts do not seem to be aware of this fact. Another key question is how models with contemporaneous and reciprocal effects fit the data relative to models without such effects. Several models are in fact equivalent, that is, they have the same number of parameters and model fit. They do, however, result in different substantive conclusions.

This paper considers identification, estimation, and model fit issues for models with contemporaneous and reciprocal effects. It explores how well the models work in practice using Monte Carlo studies as well as real-data examples. Furthermore, by using models that allow contemporaneous and reciprocal effects, the paper raises a fundamental question about current practice for cross-lagged panel modeling using models such as CLPM or RI-CLPM: Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The paper concludes that the answer is no, a finding that has important ramifications for current practice. The paper also suggests that analysts should use additional models to probe the temporalities of the CLPM, RI-CLPM effects to see if these could be considered contemporaneous rather than lagged.

The topic of reciprocal cross-lagged panel data modeling was studied over 40 years ago by Greenberg and Kessler (1982); see also Greenberg and Kessler (1979) and Kessler and Greeenberg (1981). Greenberg and Kessler (1982) demonstrated that identification can be achieved by imposing a certain degree of time invariance of the model parameters. The article, however, presented somewhat negative conclusions such as "These results are discouraging" and "the approach can be used in practice under a very restricted set of circumstances" (p. 448). Perhaps due to this, their model has not been used in recent times as far as we know. The exception is Ormel et al. (2002) who 20 years later presented an analysis using a cross-lagged model with reciprocal effects. The model also included random intercepts in line with current interest in RI-CLPM. The Greenberg-Kessler (1982) article was probably not known to the authors and was not referenced. The Ormel et al. (2002) article did not give a proof of identification but presented the claim "The full model is identified. Very different starting values gave the same solution" (p. 341). The Ormel et al. (2002) article has also not reached the audience of analysts working with cross-lagged panel data modeling perhaps due to being published in the specialized area of gerontology. In this paper, we attempt to remedy this lack of applications for panel modeling with reciprocal effects.

Section 2 sets the stage by discussing equivalent panel data models with and without contemporaneous and reciprocal effects. Section 3 discusses technical aspects of the reciprocal cross-lagged panel model for two variables including identification and estimation issues beyond those discussed in Greenberg and Kessler (1982). Section 4 presents Monte Carlo simulations using reciprocal cross-lagged panel models for different number of time points and sample sizes. Section 5 shows analyses of 5 different data sets from the literature, comparing regular RI-CLPM with reciprocal cross-lagged panel models. Section 6 concludes.

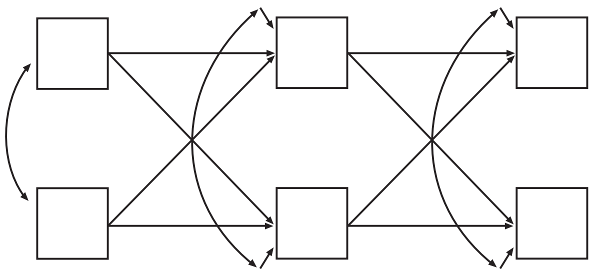
2 Equivalent models

Model equivalence is a key problem when analyzing panel data. Figure 1 shows six models for T = 3. It appears to be little known among CLPM and RI-CLPM analysts that these six models are all identified and equivalent. They are equivalent in that they have the same number of parameters and the same fit to the data. It should be noted that the same model identification and model equivalence hold when adding the random intercepts of RI-CLPM, but here the focus is on the simpler CLPM. Model (a) is the conventional CLPM with lag1 cross-lagged effects. Model (b) has no cross-lagged effects but has reciprocal effects (also referred to as a nonrecursive model; see, e.g., Bollen, 1989). It is identified by the classic econometric rule that each dependent variable has its own predictor (see, e.g., Greene, 1951, p. 325), which is due to the absence of cross-lagged effects. Model (c) includes both cross-lagged and reciprocal effects with time-invariance for the reciprocal effects, but has no residual covariances. Model (d) combines features of models (a) and (c), allowing cross-lagged effects, reciprocal effects, as well as residual covariances while imposing time invariance for both cross-lagged and reciprocal effects. Model (e) has cross-lagged effects and a contemporaneous (lag0) effect in one direction but no residual covariances. Model (f) is the same as model (e), except the lag0 effect is in the opposite direction. Because the six models have the same number of parameters and the same fit to the data, they cannot be statistically distinguished. This means, for example, that finding cross-lagged effects when using model (a) does not rule out reciprocal effects of models (b), (c), and (d), and finding reciprocal effects when using models (b), (c), and (d) does not rule out models (a), (e), and (f) with no reciprocal effects.

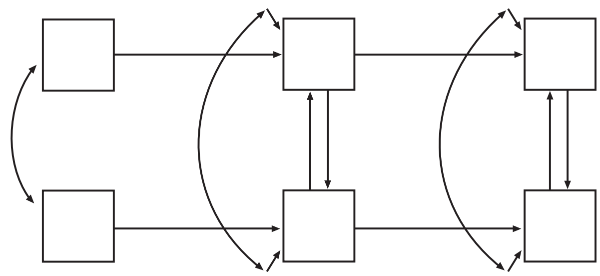
It is instructive to compare the assumptions of the regular cross-lagged model (a) and the reciprocal model (c). An advantage of model (a) is that the residual covariances allow time-varying unmeasured common causes to influence the two outcomes. In contrast, the residual covariances are assumed to be zero in model (c). Zero residual covariance may be a realistic approximation if much of the residual covariance is due to omitted contemporaneous effects. In line with regular regression, model (a) needs to assume that the residuals are uncorrelated with the two predictors, that is, the two outcomes at the previous time point. If this is not the case, the cross-lagged effects are biased. If the data have been generated by model (c), the model (a) residuals are, contrary to assumption, correlated with the predictors because each outcome at time t is influenced also by the other outcome at time t, thereby causing bias. In this way, both models make assumptions that may not be met and therefore each model has pros and cons.

Models (e) and (f) imply an additional model equivalence in that the two lag0 directions have the same number of parameters and model fit so that the direction of the contemporaneous effect cannot be statistically determined. As will be seen,

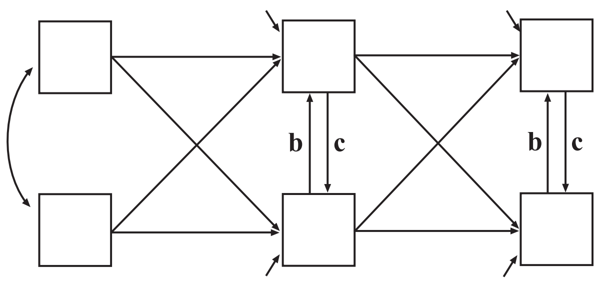
Figure 1: Six equivalent panel models for T = 3



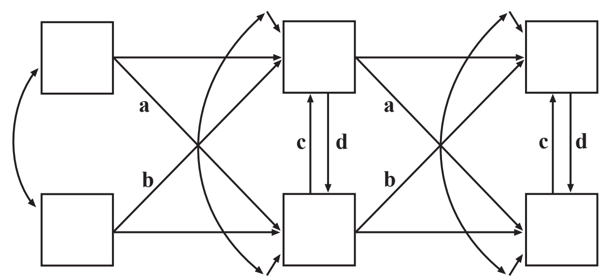
(a) CLPM: Lag1 cross-lags, residual covariances



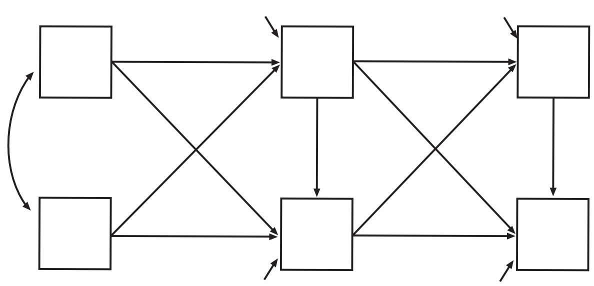
(b) RPM: Reciprocal lag0, no cross-lags, residual covariances



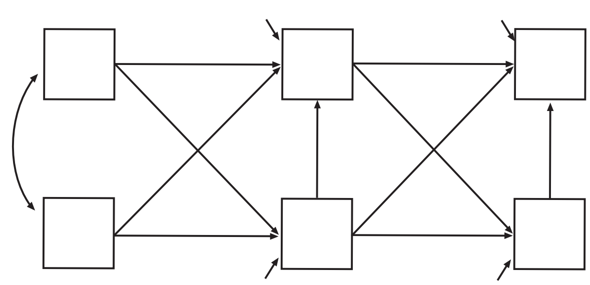
(c) RCLPM: Reciprocal lag0, lag1 cross-lags, no residual covariances



(d) IRCLPM: Invariant reciprocal lag0, lag1 cross-lags, residual covariances



(e) CLPM-Lag0: Single-direction lag0 (down), lag1 cross-lags, no residual covariances



(f) CLPM-Lag0: Single-direction lag0 (up), lag1 cross-lags, no residual covariances however, an interesting feature is that model (c) can provide supporting information to make this choice. In this sense, model (c) can be seen as a stepping stone towards more parsimonious models.

3 Modeling with reciprocal effects

This section considers technical aspects of five model variations. The first variation is the reciprocal cross-lagged model of Figure 1 (c). The second variation considers model (c) with added time invariance for not only the reciprocal effects but also the cross-lagged effects, still not including residual covariances. The third variation is model (d) with time-invariant reciprocal and cross-lagged model and added residual covariances. The fourth variation is the reciprocal model (b), which has no cross-lagged effects but includes residual covariances. The fifth variation is the contemporaneous, single-direction lag0 models (e) and (f).

3.1 RCLPM, model (c): The reciprocal cross-lagged model

Greenberg and Kessler (1982) showed identification of the reciprocal cross-lagged model in Figure 1 (c) in terms of the covariance matrix of the six variables. Here, identification of the model is instead shown by demonstrating that it is a special case of the CLPM model in Figure 1 (a). This derivation brings up issues of dual solutions and inadmissible solutions. Note that in what follows, random intercepts are not considered such as in RI-CLPM. However, the discussion below is intended to apply also for the models with random intercepts. Essentially, the focus is on the identifiability issues of the within-level model. The between-level model is standard, i.e., it would use correlated subject-specific random intercepts for all (both) variables. The random intercept / the between part of the models does not affect the within-level identifiability issues discussed below. The model and the identification issues are described below with identification proof provided in Section 1 of the Supplementary material. Readers who are less interested in the technical aspects can go straight to the summary in Section 3.1.1.

Consider the CLPM model (a) of Figure 1 for the variables Y_t and Z_t for $t = 2, ..., T$,

$$Y_t = \alpha_{yt} + \beta_{1t} Y_{t-1} + \beta_{2t} Z_{t-1} + \varepsilon_{yt} \tag{1}$$

$$Z_t = \alpha_{zt} + \beta_{3t} Y_{t-1} + \beta_{4t} Z_{t-1} + \varepsilon_{zt} \tag{2}$$

$$\varepsilon_{yt} \sim N(0, v_{yt}) \tag{3}$$

$$\varepsilon_{zt} \sim N(0, v_{zt}) \tag{4}$$

$$c_t = Cov(\varepsilon_{yt}, \varepsilon_{zt}). \tag{5}$$

Next we consider the reciprocal cross-lagged model of Figure 1 (c). This model will be referred to as RCLPM (reciprocal cross-lagged panel model). The RCLPM can be expressed as

$$Y_t = a_{yt} + r_{yt} Z_t + b_{1t} Y_{t-1} + b_{2t} Z_{t-1} + \varepsilon_{yt} \tag{6}$$

$$Z_t = a_{zt} + r_{zt} Y_t + b_{3t} Y_{t-1} + b_{4t} Z_{t-1} + \varepsilon_{zt} \tag{7}$$

$$\varepsilon_{yt} \sim N(0, w_{yt}) \tag{8}$$

$$\varepsilon_{zt} \sim N(0, w_{zt}) \tag{9}$$

$$0 = Cov(\varepsilon_{yt}, \varepsilon_{zt}). \tag{10}$$

The model is reciprocal because Y_t affects Z_t and Z_t affects Y_t. Such models are also referred to as nonrecursive models, see Bollen (1989). At time $t = 1$, both of the above models have an unrestricted model for Y_1 and Z_1 or alternatively the model is conditional on Y_1 and Z_1 and there is no distributional assumption for these variables. First note that the RCLPM has $T-1$ more parameters than the CLPM. Also, there is a simple transformation that converts the RCLPM model into the CLPM. It is easier to illustrate the transformation with matrix notation.

The CLPM in matrix form is

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} \alpha_{yt} \\ \alpha_{zt} \end{pmatrix} + \begin{pmatrix} \beta_{1t} & \beta_{2t} \\ \beta_{3t} & \beta_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} \tag{11}$$

where

$$Var \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} v_{yt} & c_t \\ c_t & v_{zt} \end{pmatrix}.$$

The RCLPM in matrix form is

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} + \begin{pmatrix} r_{yt} & 0 \\ 0 & r_{zt} \end{pmatrix} \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} + \begin{pmatrix} b_{1t} & b_{2t} \\ b_{3t} & b_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} \tag{12}$$

where

$$Var \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} w_{yt} & 0 \\ 0 & w_{zt} \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} + \begin{pmatrix} b_{1t} & b_{2t} \\ b_{3t} & b_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}. \tag{13}$$

Since

$$\begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix}^{-1} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix},$$

the RCLPM becomes

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} + \tag{14}$$

$$\frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} b_{1t} & b_{2t} \\ b_{3t} & b_{4t} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \tag{15}$$

$$\frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}. \tag{16}$$

The above equation is a structured CLPM. In the CLPM (11), all parameters are unrestricted. The above expression of the RCLPM reveals that it is a CLPM where the parameters have certain structural form. This implies that the RCLPM is nested within the CLPM and the CLPM parameters can be obtained from the RCLPM parameters as follows

$$\begin{pmatrix} \alpha_{yt} \\ \alpha_{zt} \end{pmatrix} = \frac{1}{1 - r_{yt} r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} \tag{17}$$

$$\begin{pmatrix} \beta_{1t} & \beta_{2t} \\ \beta_{3t} & \beta_{4t} \end{pmatrix} = \frac{1}{1 - r_{yt}r_{zt}} \begin{pmatrix} 1 & r_{yt} \\ r_{zt} & 1 \end{pmatrix} \begin{pmatrix} b_{1t} & b_{2t} \\ b_{3t} & b_{4t} \end{pmatrix} \tag{18}$$

$$v_{yt} = \frac{w_{yt} + r_{yt}^2 w_{zt}}{(1 - r_{yt}r_{zt})^2} \tag{19}$$

$$v_{zt} = \frac{w_{zt} + r_{zt}^2 w_{yt}}{(1 - r_{yt}r_{zt})^2} \tag{20}$$

$$c_t = \frac{w_{zt}r_{yt} + w_{yt}r_{zt}}{(1 - r_{yt}r_{zt})^2}. \tag{21}$$

Note that equations (17-18) are completely reversible (one-to-one transformation), given particular values of r_{yt} and r_{zt} as long as $1 \neq r_{yt}r_{zt}$. This means that given the r_{yt} and r_{zt} values, the parameters of the RCLPM can be obtained from the parameters of the CLPM

$$\begin{pmatrix} a_{yt} \\ a_{zt} \end{pmatrix} = \begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \alpha_{yt} \\ \alpha_{zt} \end{pmatrix} \tag{22}$$

$$\begin{pmatrix} b_{1t} & b_{2t} \\ b_{3t} & b_{4t} \end{pmatrix} = \begin{pmatrix} 1 & -r_{yt} \\ -r_{zt} & 1 \end{pmatrix} \begin{pmatrix} \beta_{1t} & \beta_{2t} \\ \beta_{3t} & \beta_{4t} \end{pmatrix}. \tag{23}$$

Therefore, the relationship between the two models is entirely dependent on equations (19-21). In these equations, the CLPM has $3(T-1)$ parameters v_{yt}, v_{zt}, c_t, for $t = 2, ..., T$, while the RCLPM has $4(T-1)$, r_{yt}, r_{zt}, w_{yt}, w_{zt}. Once again we see here that the RCLPM has additional $T-1$ parameters but is nevertheless nested within the CLPM. Therefore, for the RCLPM to be identified, it must have at least $T-1$ additional parameter constraints. If we introduce exactly $T-1$ constraints, it is possible to produce an RCLPM that is equivalent to the CLPM. The proof of this is given in Section 1 of the Supplementary material.

For $T = 3$, the RCLPM is equivalent to the CLPM if we constrain the reciprocal interactions to be invariant across time. Similarly, we can show that for $T = 2n+1$, the RCLPM is equivalent to the CLPM if we constrain neighboring reciprocal interactions to be the same. That is, if we constrain the reciprocal interaction for times 2 and 3 to be the same, and the reciprocal interaction for times 4 and 5 to be the same etc..., the model becomes equivalent to the CLPM. From this we can further conclude that the RCLPM with time invariant reciprocal interactions across all time points is identified and is nested within the CLPM.

It should be emphasized that the identification of the RCLPM requires a sufficient amount of time non-invariance. With time series data, the RCLPM with complete time invariance across all parameters is a two-level vector auto-regressive dynamic structural equation (DSEM-VAR) model discussed in Asparouhov et al. (2018). It is well known that this is not an identified model. For RCLPM, certain data sets may result in a high negative correlation between the two reciprocal effect estimates when auto-regressions and/or residual variances do not have sufficient variation across time even though the model allows them to vary across time. This is seen in the Section 3.2 equations (24), (25) and discussed below these equations. It is also seen in equations (26) - (30) of Section 1 in the Supplementary material. When this correlation is -1, it can be considered an empirical nonidentification.[1]

[1] In Mplus, the size of the correlation between the two reciprocal effect estimates is shown in TECH3.

3.1.1 Summary and guidelines

The previous section discussed the identification of RCLPM when T - 1 time invariance restrictions are imposed on the reciprocal effects. The identification proof in Section 1 of the Supplementary material demonstrates that the RCLPM has a dual solution issue. Sections 2 and 3 of the Supplementary material discuss ways to resolve the dual solution issue using the reciprocal effect constraint $(r_y r_z)^2 < 1$, referred to as restriction (b). To avoid both dual solutions and negative R^2 solutions, the reciprocal effects constraint $0 < r_y r_z < 1$ is used, referred to as restriction (a). Following are general guidelines for analysis using the RCLPM.

1. Estimate the RCLPM with the reciprocal effects constraint $0 < r_y r_z < 1$ to avoid duality and negative R^2 solution. Random starting values should be used in this estimation.[2]
2. If both reciprocal regressions parameters are significant, the RCLPM can be considered fully interpretable and supported by the data.
3. If one or both reciprocal regressions parameters are not significant and are not zero, these parameters can be eliminated from the model. The RCLPM can be converted to a much simpler model without reciprocal regressions.
4. If one of the reciprocal regression parameters is estimated as 0, this means that the estimation terminated at the border of the allowed parameter space. This can be taken as evidence that both reciprocal effects are not supported. In this case, the recommendation is to fix the parameter at this 0 boundary value. This would guarantee correct standard errors for the rest of the parameters and result in the same loglikelihood value.

Under some circumstances, it may be necessary to pursue a reciprocal model even when the parameters are not significant. Reciprocal regression parameters tend to have larger standard errors and establishing significance may require a large sample size which may not be available. If the reciprocal regression parameters appear to be substantial in a standardized metric, the RCLPM could be used as an exploration even if one or both of the reciprocal regressions are not significant.

3.2 RCLPM with time-invariant reciprocal and cross-lagged regressions

This section considers the second reciprocal model variation of RCLPM with time-invariant reciprocal and cross-lagged regressions. It is shown that this model does not have a dual solution, i.e., introducing the constraint of invariant cross-lagged regressions is sufficient to eliminate the dual solution. For this model, the cross-lagged relations b_{2t} and b_{3t} as well as the reciprocal regressions r_y and r_z are time invariant. In this case, the dual solution is removed as long as the auto-regressive parameters b_{1t} and b_{4t} are not time invariant. Using equation (18), we conclude that β_{1t} and β_{4t} are also not time invariant. Then using equation (23) for $t = 2$ and $t = 3$, we get that

$$\beta_{22} - r_y \beta_{42} = b_{22} = b_{23} = \beta_{23} - r_y \beta_{43}$$

[2]This can be done via the STARTS option in Mplus.

and therefore

$$r_y = \frac{\beta_{22} - \beta_{23}}{\beta_{42} - \beta_{43}} \tag{24}$$

which yields a unique solution. Similarly

$$r_z = \frac{\beta_{32} - \beta_{33}}{\beta_{12} - \beta_{13}}. \tag{25}$$

We used information from time points $t = 2$ and $t = 3$ only but any other time points will yield the same conclusion.

It should be noted here that the stability of the reciprocal regression estimates is very dependent on the non-invariance of the auto-regressive parameters b_{1t} and b_{4t}, which is tightly connected to the non-invariance of β_{1t} and β_{4t}. This is also discussed in Greenberg and Kessler (1982) in terms of non-identification if the process is in equilibrium, that is, fully time invariant. If the non-invariance is weak and the distribution of the denominators in the above formulas approach zero, the reciprocal regression parameters may be somewhat poorly identified, may exhibit a dual solution problem, may have large standard errors and confidence intervals, and may exhibit highly skewed parameter distributions. In such a case, using the reciprocal effect constraint of $(r_y r_z)^2 < 1$, bootstrap and Bayesian estimation methods are preferred as these can accommodate skewed parameter distributions and provide more accurate non-symmetric confidence intervals.

Note that RCLPM with invariant reciprocal and cross-lagged regressions does not avoid the possible negative R^2 issue discussed in Section 3 of the Supplementary material.

3.3 IRCLPM, model (d): RCLPM with time-invariant reciprocal and cross-lagged regressions and non-invariant residual covariances

This section considers the third model variation of RCLPM where non-invariant residual covariances are added to the model with invariant reciprocal and cross-lagged regressions. This model will be referred to as IRCLPM. Section 4 of the Supplementary material shows that time-specific residual covariances can be added to the RCLPM as long as the reciprocal and the cross-lagged parameters are held time invariant. It is also necessary that the auto-regressive parameters b_{1t} and b_{4t} are not time invariant which holds when the auto-regressive parameters β_{1t} and β_{4t} are not time invariant. It is possible to further constrain the residual covariance or the residual correlation to be time invariant. A sufficient condition to ensure that the solution has positive R^2 values is if the reciprocal parameters r_y, r_z and the residual covariances have the same signs and $0 < r_y r_z < 1$.

3.4 RPM, model (b): RCLPM without cross-lagged regressions

This section considers the fourth model variation of RCLPM without the cross-lagged regressions but with residual covariance. This model will be referred to as RPM. The

RPM is given by the following equations

$$Y_t = a_{yt} + r_{yt}Z_t + b_{1t}Y_{t-1} + \varepsilon_{yt}$$

$$Z_t = a_{zt} + r_{zt}Y_t + b_{4t}Z_{t-1} + \varepsilon_{zt}$$

$$\varepsilon_{yt} \sim N(0, w_{yt})$$

$$\varepsilon_{zt} \sim N(0, w_{zt})$$

$$w_t = Cov(\varepsilon_{yt}, \varepsilon_{zt}).$$

The model has the same number of parameters as the CLPM and in fact the two models are equivalent as shown in Section 5 of the Supplementary material. The identification of the RPM without cross-lags does not require equality constraints across-time, i.e., the reciprocal regression parameters can be time-specific.

Note that the absence of cross-lagged regressions in the RPM avoids the dual solution problem discussed in Section 2 of the Supplementary material but it does not avoid the possible negative R^2 issue discussed in Section 3 of the Supplementary material.

3.5 CLPM-Lag0, models (e) and (f): Contemporaneous, single-direction lag0 models

In practical applications, a common scenario will be that one of the two reciprocal regression parameters will not be significant. The question arises if instead of reciprocal modeling, there are advantages to adding a single contemporaneous regression parameter to the regular cross-lagged modeling, doing two analyses with the contemporaneous effect in opposite directions. Following is a list of such models and their identification status with CE denoting the contemporaneous effect. Parameters not mentioned are time varying.

1. Time varying CE, time varying residual covariances: Not identified
2. Time invariant CE, time varying residual covariances: Not identified
3. Time invariant CE, time invariant residual covariances: Identified unless all residual variances are also time invariant
4. Time invariant CE, time invariant cross-lags, time varying residual covariances: Identified by the fact that the IRCLPM is identified
5. Time varying CE, no residual covariances: Identified because CLPM is identified

Model variation 3 is of interest when using the version where the residual correlations are held time invariant. It will be applied to the 5 examples. Model variation 5 is the same as models (e) and (f) in Figure 1. This variation is straightforward and can be seen as a follow-up analysis when the reciprocal model finds a significant lag0 effect in only one direction.

4 Monte Carlo simulations

Because little is known about how well analysis with the reciprocal cross-lagged model works in practice, it is of interest to study how it performs in Monte Carlo simulation studies. Throughout, the random intercept version of the models is used. The random intercept version of the model in Figure 1 (c), RI-RCLPM, is studied first, followed by the random intercept version of Figure 1 (d) model with invariant reciprocal and cross-lagged effects and with residual covariances, referred to as RI-IRCLPM where the added I stands for invariant. The aims are to determine how well parameter values can be recovered, the quality of standard error estimation, the coverage, and the power to detect reciprocal effects. The special case of RI-RCLPM where one of the reciprocal effects is zero is also studied. It is shown to provide a way to determine the direction of the single-direction lag0 model, that is, the random intercept counterpart to models (e) and (f) in Figure 1.

4.1 Performance of the RI-RCLPM

Parameter values for the RI-RCLPM are based on example 1 (MWI) to be discussed in Section 5. Here, T = 5. The reciprocal effects are time invariant in both data generation and analyses. The cross-lagged effects are time invariant in the data generation but time invariance is not imposed in the analysis. The example features one reciprocal effect that is medium-sized (-0.4) in a standardized metric and one cross-lagged effect that is small (-0.1) in a standardized metric. For simplicity, missing data due to attrition is not represented in the data generation. Multivariate normal, continuous variables are generated using 500 replications. For each replication, 500 bootstrap draws are made to compute standard errors and to capture non-normal parameter estimate distributions and create non-symmetric confidence intervals. In addition, robust maximum-likelihood standard errors and symmetric confidence intervals are computed using the Mplus option MLR. The non-duality restriction (b), $(r_y r_z)^2 < 1$, is imposed on the reciprocal effects.[3] The number of time points T is varied as 3, 4, 5 where T = 3 and 4 runs are based on real-data estimates for the first 3 and 4 time points. Sample size N is varied as 500, 750, and 1000. The 5% χ^2 reject proportion for the replications is close to the correct value of 0.05 and is not reported.

Example 1 that the simulation study builds on considers the relationship between the two variables Self-esteem and Depression, referred to as S and D in the following. Figure 2 and Figure 3 show the distribution of the reciprocal estimate $S_t \rightarrow D_t$ over the Monte Carlo replications for T = 5 and T=3, respectively. As expected, the figures show a slight non-normality but the skewness is only 0.107 for T = 5 and 0.513 for T = 3.[4] Because the non-normality is not pronounced, Monte Carlo results will be presented using both bootstrap and MLR.

Table 1 shows the results for the three sample sizes with T = 5. The first column shows the auto-regressive, cross-lagged, and reciprocal parameters at time points 4 and 5 for the S variable and at time points 1 and 2 for the D variable (the hat notation refers to the within-level version of the variable). The second and third columns shows the

[3] Restriction (a) may be needed for the bootstrap runs.

[4] The plots are obtained by using the Mplus RESULTS option of the MONTECARLO command to save estimates for all replications, followed by a TYPE=BASIC run on the saved file to plot the distribution.

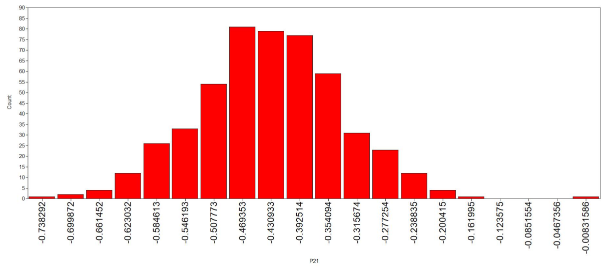


Figure 2: Monte Carlo distribution of the reciprocal effect $S_t \rightarrow D_t$ for N = 500 and T = 5

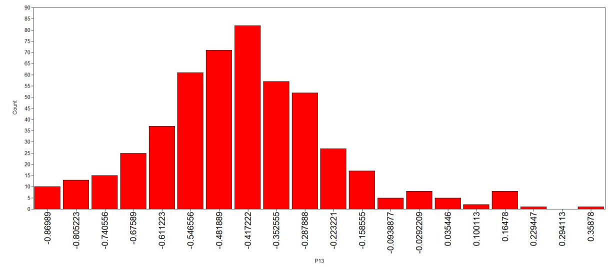


Figure 3: Monte Carlo distribution of the reciprocal effect $S_t \rightarrow D_t$ for N = 500 and T = 3

parameter values generating the data and the average estimates over the replications. The fourth column shows the standard deviation over the replications which is used for comparison with the fifth and sixth columns of standard error averages over the replications using bootstrap and MLR standard errors, respectively. The 7th and 8th columns show the mean squared error (M.S.E.) of the estimate and the 95% coverage using bootstrap. The last 2 columns show the power to reject a zero parameter value as judged by the proportion of replications for which the confidence interval does not include zero using the non-symmetric bootstrap confidence interval and the symmetric MLR confidence interval, respectively.

Table 1 shows that the parameter values are well recovered for all 3 sample sizes. The bootstrap standard errors are on the whole a bit overestimated while the MLR standard errors perform very well. The 95% coverage is good overall. Of key interest is the power to reject that the effect $S_t \rightarrow D_t$ is zero. This is the parameter labeled D2ˆ ON S2ˆ in the first column. The population value for this effect is -0.431 which corresponds to a medium-sized standardized effect. For N = 1000, the desired power of 0.80 is reached as estimated by both bootstrap and MLR. For N = 500, the bootstrap estimate is only 0.564 while the MLR estimate is more optimistic. The lower power estimate for bootstrap may be due to the overestimated standard error, resulting in a wider confidence interval.

Table 2 and Table 3 show the corresponding results for T = 4 and T = 3, respectively. As expected, the power diminishes with fewer time points and smaller sample sizes. With N = 500, it is clear that the power is too low to reject a zero effect for $S_t \rightarrow D_t$ when T = 3 and T = 4.

4.1.1 The importance of time varying auto-regression coefficients and residual variances for RI-RCLPM

Section 3 pointed to the importance of variation across time in parameter values for the RI-RCLPM. If there is time invariance of all cross-lagged, reciprocal, and auto-regression parameters, the model is not identified. If there is variation but it is not large, large standard errors result. Of particular importance is variation across time in auto-regression coefficients as discussed in connection with equations (24) and (25). If this variation is small, standard errors can get so large that the model becomes useless. In the simulations just discussed, there is rather small variation in the auto-

Table 1: Monte Carlo results for RI-RCLPM, T = 5

	ESTIMATES Population	Average	Std. Dev.	S. E. Average	**MLR** S.E. Average	M. S. E.	95% Cover	% Sig Coeff	% Sig **MLR**
				T = 5, N = 1000					
S5ˆ ON									
S4ˆ	0.510	0.5121	0.0455	0.0550	0.0491	0.0021	0.976	1.000	1.000
D4ˆ	-0.158	-0.1630	0.0674	0.0835	0.0729	0.0046	0.964	0.670	0.680
D5ˆ	-0.091	-0.0760	0.1680	0.2242	0.1789	0.0284	0.964	0.060	0.120
D2ˆ ON									
D1ˆ	0.214	0.2135	0.0417	0.0469	0.0438	0.0017	0.962	0.992	0.990
S1ˆ	0.059	0.0569	0.0381	0.0394	0.0378	0.0014	0.946	0.294	0.334
S2ˆ	-0.431	-0.4323	0.0966	0.1251	0.1028	0.0093	0.970	0.844	**0.954**
				T = 5, N = 750					
S5ˆ ON									
S4ˆ	0.510	0.5114	0.0543	0.0665	0.0612	0.0029	0.972	1.000	0.992
D4ˆ	-0.158	-0.1667	0.0894	0.0994	0.0912	0.0080	0.962	0.504	0.480
D5ˆ	-0.091	-0.0693	0.2267	0.2704	0.2319	0.0518	0.962	0.048	0.112
D2ˆ ON									
D1ˆ	0.214	0.2141	0.0505	0.0560	0.0631	0.0025	0.960	0.966	0.964
S1ˆ	0.059	0.0582	0.0427	0.0472	0.0445	0.0018	0.964	0.228	0.254
S2ˆ	-0.431	-0.4311	0.1265	0.1535	0.1303	0.0160	0.962	0.712	**0.876**
				T = 5, N = 500					
S5ˆ ON									
S4ˆ	0.510	0.5115	0.0745	0.0830	0.0782	0.0055	0.970	1.000	0.992
D4ˆ	-0.158	-0.1693	0.1136	0.1211	0.1171	0.0130	0.956	0.368	0.354
D5ˆ	-0.091	-0.0644	0.2925	0.3214	0.2989	0.0861	0.968	0.056	0.130
D2ˆ ON									
D1ˆ	0.214	0.2157	0.0629	0.0699	0.0672	0.0040	0.964	0.898	0.898
S1ˆ	0.059	0.0551	0.0546	0.0610	0.0566	0.0030	0.974	0.118	0.166
S2ˆ	-0.431	-0.4255	0.1639	0.1855	0.1681	0.0269	0.964	0.564	**0.792**

Table 2: Monte Carlo results for RI-RCLPM, T = 4

	ESTIMATES Population	Average	Std. Dev.	S. E. Average	MLR S.E. Average	M. S. E.	95% Cover	% Sig Coeff	% Sig MLR
T = 4, N = 1000									
S4^ ON									
S3^	0.412	0.4113	0.0503	0.0574	0.0538	0.0025	0.974	1.000	0.998
D3^	-0.158	-0.1652	0.0911	0.1004	0.0911	0.0083	0.964	0.451	0.431
D4^	-0.091	-0.0642	0.1969	0.2371	0.2043	0.0394	0.966	0.048	0.100
D2^ ON									
D1^	0.214	0.2142	0.0495	0.0515	0.0488	0.0024	0.956	0.988	0.982
S1^	0.059	0.0591	0.0419	0.0434	0.0415	0.0018	0.950	0.248	0.273
S2^	-0.431	-0.4369	0.1073	0.1299	0.1122	0.0115	0.962	0.830	**0.934**
T = 4, N = 750									
S4^ ON									
S3^	0.412	0.4103	0.0604	0.0684	0.643	0.0036	0.968	1.000	1.000
D3^	-0.158	-0.1688	0.1064	0.1171	0.1109	0.0116	0.966	0.357	0.301
D4^	-0.091	-0.0577	0.2466	0.2765	0.2540	0.0618	0.978	0.040	0.116
D2^ ON									
D1^	0.214	0.2141	0.0578	0.0610	0.0587	0.0033	0.970	0.964	0.956
S1^	0.059	0.0595	0.0481	0.0519	0.0492	0.0023	0.962	0.210	0.240
S2^	-0.431	-0.4348	0.1325	0.1523	0.1362	0.0175	0.972	0.727	**0.866**
T = 4, N = 500									
S4^ ON									
S3^	0.412	0.4079	0.0773	0.0876	0.0845	0.0060	0.972	0.988	0.974
D3^	-0.158	-0.1763	0.1464	0.1438	0.1499	0.0217	0.956	0.251	0.196
D4^	-0.091	-0.0427	0.3349	0.3324	0.3491	0.1143	0.956	0.059	0.119
D2^ ON									
D1^	0.214	0.2126	0.0729	0.0765	0.0777	0.0053	0.962	0.834	0.826
S1^	0.059	0.0581	0.0561	0.0669	0.0625	0.0031	0.976	0.115	0.145
S2^	-0.431	-0.4329	0.1774	0.1857	0.1911	0.0314	0.960	0.600	**0.739**

Table 3: Monte Carlo results for RI-RCLPM, T = 3

	ESTIMATES Population	Average	Std. Dev.	S. E. Average	MLR S.E. Average	M. S. E.	95% Cover	% Sig Coeff	% Sig MLR
T = 3, N = 1000									
S3ˆ ON									
S2ˆ	0.020	0.0195	0.1141	0.1237	0.1183	0.0130	0.956	0.040	0.036
D2ˆ	-0.158	-0.1623	0.0875	0.1020	0.0961	0.0077	0.968	0.396	0.420
D3ˆ	-0.091	-0.0452	0.3015	0.2900	0.2912	0.0928	0.940	0.068	0.080
D2ˆ ON									
D1ˆ	0.214	0.2182	0.0602	0.0663	0.0666	0.0036	0.954	0.944	0.918
S1ˆ	0.059	0.0594	0.0621	0.0598	0.0583	0.0039	0.946	0.180	0.178
S2ˆ	-0.431	-0.4339	0.1337	0.1414	0.1401	0.0179	0.948	0.784	**0.860**
T = 3, N = 750									
S3ˆ ON									
S2ˆ	0.020	0.0176	0.1310	0.1486	0.1406	0.0171	0.958	0.036	0.034
D2ˆ	-0.158	-0.1587	0.1094	0.1232	0.1183	0.0119	0.964	0.282	0.276
D3ˆ	-0.091	-0.0221	0.3642	0.3221	0.3427	0.1372	0.938	0.068	0.093
D2ˆ ON									
D1ˆ	0.214	0.2189	0.0699	0.0764	0.0771	0.0049	0.958	0.875	0.827
S1ˆ	0.059	0.0631	0.0704	0.0708	0.0687	0.0050	0.948	0.145	0.169
S2ˆ	-0.431	-0.4346	0.1662	0.1597	0.1632	0.0276	0.934	0.710	**0.791**
T = 3, N = 500									
S3ˆ ON									
S2ˆ	0.020	0.0105	0.1666	0.1965	0.1788	0.0278	0.974	0.022	0.022
D2ˆ	-0.158	-0.1639	0.1338	0.1653	0.1540	0.0179	0.976	0.163	0.187
D3ˆ	-0.091	-0.0054	0.4363	0.3743	0.4169	0.1974	0.936	0.070	0.099
D2ˆ ON									
D1ˆ	0.214	0.2260	0.0946	0.0961	0.0967	0.0091	0.960	0.708	0.648
S1ˆ	0.059	0.0618	0.0862	0.0908	0.0876	0.0074	0.952	0.109	0.113
S2ˆ	-0.431	-0.4311	0.2003	0.1948	0.2021	0.0400	0.940	0.584	**0.682**

regression coefficients. This is due to the simulation being built on example 1 (MWI) to be discussed in Section 5, where especially the depression outcome has very similar auto-regression coefficients over time. While this small variation did not harm the simulations presented, changing the smaller reciprocal parameter to a larger value in the simulation made the analysis fail with large standard errors and parameter bias. In contrast, basing the simulation on example 3 (BLS), the larger variation in auto-regressions made it possible to obtain good results even when the reciprocal effects were set to be large and of equal size. In practical terms, real-data analysis using RI-RCLPM should not impose time-invariance of auto-regressions.

Section 3.1 discussed the potential of a high negative correlation between the two reciprocal effect estimates due to data with little time variation in auto-regressions and residual variances. For the Table 1 simulation with T = 5, N = 1000, the correlations for the first 5 replications range from -0.976 to -0.987, indicating that the model is empirically weakly identified. It is interesting that the estimation of the model still performs well. The sensitivity to time variation in the data is illustrated by the fact that changing just one of the residual variances to a higher population value in the data generation causes a substantial drop in this correlation.

4.2 Performance of the RI-IRCLPM

Parameter values for the Section 3.3 model RI-IRCLPM that imposes time invariance of both cross-lagged and reciprocal effects and adds residual covariances are based on example 4 (NLSY) to be discussed in Section 5. The example has T = 11, but here only the first five time points are used so that T = 5 as for the RI-RCLPM simulation. As in Section 5, a version of the model is used that imposes time invariance of the residual correlations. This example has one small-sized reciprocal effect (-0.15). Due to the small effect size, this example considers T = 2000, T = 1000, and T = 500.

Table 4 shows the results using the MLR estimator. The parameter values are well recovered for all 3 sample sizes. The standard errors are well estimated and the 95% coverage is good overall. Of key interest is the power to reject that the effect $D_t \rightarrow S_t$ is zero. This is the parameter labeled S5ˆ ON D5ˆ in the first column. The population value for this effect is -0.104 which corresponds to a small-sized standardized effect of -0.15. For N = 2000, the power is estimated as 0.96, exceeding the desired power of 0.80. For N = 1000, the power drops to 0.76 and for N = 500, the power is only 0.46.

4.3 Using RI-RCLPM to determine the direction of single-direction lag0 modeling

A special case of RI-RCLPM is when one of the reciprocal effects is zero. Changing the smaller reciprocal effect $D_t \rightarrow S_t$ of -0.091 to zero in the Section 4.1 data generation while estimating both effects using RI-RCLPM produces good simulation result. This suggests that RI-RCLPM can be used to determine the direction of single-direction lag0 modeling, that is, the random intercept counterpart to the model of Figure 1 (d). To illustrate this, 500 replications of N = 1000, T = 5 data were generated using the RI-RCLPM with the time-invariant $D_t \rightarrow S_t$ effect set to zero. Based on these data, two analyses were carried out using the single-direction, non-invariant lag0 effects model, one analysis for each direction. Table 5 shows the average estimates over the 500

Table 4: Monte Carlo results for RI-IRCLPM, T = 5

	ESTIMATES			S. E.	M. S. E.	95%	% Sig
	Population	Average	Std. Dev.	Average		Cover	Coeff
			T = 5, N = 2000				
S5ˆ ON							
S4ˆ	-0.173	-0.1741	0.0321	0.0324	0.0010	0.946	1.000
D4ˆ	-0.005	-0.0052	0.0098	0.0097	0.0001	0.960	0.088
D5ˆ	-0.104	-0.1045	0.0277	0.0273	0.0008	0.940	**0.962**
D2ˆ ON							
D1ˆ	0.161	0.1612	0.0241	0.0241	0.0006	0.956	1.000
S1ˆ	-0.043	-0.0416	0.0238	0.0233	0.0006	0.946	0.424
S2ˆ	-0.018	-0.0216	0.0568	0.0559	0.0032	0.934	0.078
			T = 5, N = 1000				
S5ˆ ON							
S4ˆ	-0.173	-0.1727	0.0427	0.0460	0.0018	0.960	0.978
D4ˆ	-0.005	-0.0053	0.0140	0.0137	0.0002	0.942	0.068
D5ˆ	-0.104	-0.1046	0.0391	0.0391	0.0015	0.948	**0.760**
D2ˆ ON							
D1ˆ	0.161	0.1584	0.0345	0.0342	0.0012	0.948	0.998
S1ˆ	-0.043	-0.0414	0.0341	0.0331	0.0012	0.934	0.244
S2ˆ	-0.018	-0.0197	0.0811	0.0801	0.0066	0.948	0.060
			T = 5, N = 500				
S5ˆ ON							
S4ˆ	-0.173	-0.1743	0.0640	0.0653	0.0041	0.962	0.778
D4ˆ	-0.005	-0.0063	0.0201	0.0196	0.0004	0.950	0.066
D5ˆ	-0.104	-0.1028	0.0555	0.0566	0.0031	0.958	**0.458**
D2ˆ ON							
D1ˆ	0.161	0.1581	0.0473	0.0487	0.0022	0.948	0.902
S1ˆ	-0.043	-0.0459	0.0484	0.0472	0.0023	0.952	0.172
S2ˆ	-0.018	-0.0128	0.1143	0.1158	0.0131	0.966	0.040

Table 5: Two single-direction lag0 analyses using RI-RCLPM data

$S_t \rightarrow D_t$ (true value = -0.431)			
t=2	t=3	t=4	t=5
-0.4337 (.0743)	-0.4318 (.0293)	-0.4306 (.0285)	-0.4321 (.0281)
$D_t \rightarrow S_t$ (true value = 0)			
t=2	t=3	t=4	t=5
-0.3907 (.0704)	-0.6730 (.0476)	-0.5784 (.0386)	-0.5435 (.0357)

replications for the lag0 effects in the two directions. The top part of the table shows that the true value of -0.431 for the $S_t \rightarrow D_t$ effect is well estimated with small standard errors for each of the timepoints. The zero $D_t \rightarrow S_t$ effect, however, obtains estimates significantly different from zero. In fact, three of the four estimates are larger than for the effect in the opposite direction. Because the standard errors are small, this would result in the misleading conclusion of significant effects in the wrong direction. In line with the discussion of equivalent models in Section 2, the model fit is exactly the same for the two models so model fit cannot be used to determine direction. The fit is good because the data were generated by RI-RCLPM with one reciprocal effect being zero.

In contrast to the single-direction lag0 modeling, RI-RCLPM results for the same generated data are good. The RI-RCLPM average estimate for the $D_t \rightarrow S_t$ effect is zero and the average estimate for the $S_t \rightarrow D_t$ effect is -0.4320 with average standard error 0.1067, only a little higher than the standard errors for the single-direction model. This supports the notion that RI-RCLPM can be used to determine the direction of single-direction lag0 modeling. This will be illustrated in the analyses of the five examples discussed in Section 5.

5 Analysis of 5 examples

To illustrate the performance of the reciprocal effect modeling in real-data settings, five examples from the literature are re-analyzed, covering a wide range of sample sizes, time points, and time intervals between measurements. Three of the examples are from Orth et al. (2021) concerning depression and self-esteem: MWI, BLS, and NLSY. One example is from Ormel et al. (2002) concerning depression and disability and one example is from Nunez-Regueiro et al. (2021) concerning academic self-concept and achievement (GPA):

1. MWI data: N = 663, T = 5, Interval = 2 months
2. Ormel data: N = 753, T = 3, Interval = 1 year
3. BLS data: N = 404, T = 4, Interval = 1 year
4. NLSY data: N = 8,259, T = 11, Interval = 2 years
5. GPA data: N = 933, T = 5, Interval = 4 months

In all cases, random intercepts are used in the model for four reasons: (1) It is possible that there are individual differences in the level (average over time) of the outcomes (i.e., the data may have trait-like features), (2) the fit is better than not having them in the model, (3) the random intercept variances are substantial relative to their standard errors, and (4) the random intercepts are motivated by the statistical principles of multilevel modeling for hierarchical data with both within- and between-person variation (Hamaker et al., 2015; Hamaker, 2023).

5.1 Example 1: The MWI data on depression and self-esteem

Orth et al. (2021) studied the relationship between self-esteem and depression. They described four models: (1) the vulnerability model that postulates that low self-esteem leads to depression, (2) the scar model that postulates that low self-esteem is a consequence of depression, (3) the reciprocal relation model that allows influence in both directions, and (4) the third factor model where e.g. prior stressful life events or underlying temperament factors cause a spurious influence on both outcomes.

One of the data sets studied in Orth et al. (2021) is My Work and I (MWI) which is an adult sample of N = 663 observed over 5 time points with a time interval of 2 months. The coverage declines from 0.99 to 0.57, that is, 57% of the sample remain at the end of the 5 time points. The measurements are as follows:

- Self-esteem: Participants were asked how much they agree with each of the statements included in the scale (no time frame stated, so could include current and past status)
- Depression: Participants were instructed to assess how frequently they had experienced each symptom within the preceding 30 days

Analysis using regular RI-CLPM with time invariant cross-lagged effects points to a small but significant negative cross-lagged effect of depression on self-esteem, $D_{t\text{-}1} \rightarrow S_t$ (Orth et al., 2021, Table 6). For the reciprocal RI-CLPM (RI-RCLPM), two analyses are carried out. Analysis using the MLR estimator takes into account the non-normality of the variables in the chi-square and standard error computations but does not provide non-symmetric confidence intervals which may be needed for the reciprocal effects. Bootstrap analysis uses ML[5] and gives bootstrap standard errors and bootstrap non-symmetric confidence intervals matching a skewed distribution for the reciprocal estimates. Based on the Monte Carlo simulations, however, the bootstrap

[5] Note that in Mplus, MLR parameter estimates = ML parameter estimates = parameter estimates using bootstrap

standard errors are a bit inflated and the confidence intervals may be a bit too wide (conservative).

The Mplus input file for RI-RCLPM showing both the MLR and bootstrap analyses is presented in Figure 4. The MODEL command statements use the hat notation to denote within-level variables as presented in Asparouhov and Muthén (2022) and discussed in Mplus Web Talk No. 4, Part 1. The hat notation refers to residuals, in this case residuals in the regression of each observed variable on the random intercept. Previously, these variables had to be defined in a more cumbersome way using BY statements in line with factor analysis and adding the specification of zero measurement error. For the reciprocal effect part, time invariance is imposed using the labels (rsd) and (rds). The MODEL CONSTRAINT command shows the two alternative restrictions imposed on the reciprocal effects as discussed in Section 3 called (a) and (b) here.

To obtain the regular RI-CLPM, the Mplus input of Figure 4 should be modified by deleting the reciprocal statements and adding residual covariances for all time points. To obtain the RI-RPM, the cross-lagged statements should be deleted, the time invariance of the reciprocal effects deleted together with MODEL CONSTRAINT, and residual covariances added for all time points. To obtain the Section 3.3 model RI-IRCLPM, Figure 5 shows the new MODEL statements for residual covariances and variances and the MODEL CONSTRAINT statements needed for time invariant residual correlations as well as for imposing the constraint on the reciprocal effects that enforces the restriction discussed in Section 4 of the Supplementary material, 0 <(r1*r2) <1 and 0 >r1*rho where rho is the time-invariant residual correlation.

As a reminder of the discussion in Section 2 and Section 3, Figure 6 shows five key models for T = 5 with notation in line with Figure 1, adding the prefix RI for random intercept. Only the within-level part is drawn, not the between-level random intercept part. The circles denote the latent within part of the observed variables after subtracting the between part. Model (a) represents the regular RI-CLPM, model (b) represents the RI-RPM, model (c) represents the RI-RCLPM, model (d) represents the RI-IRCLPM, and model (e) represents the RI-CLPM-Lag0 (only the arrow down version shown). RI-CLPM and RI-RPM have the same number of parameters. This number of parameters is also obtained by imposing T - 1 = 4 restrictions on the reciprocal effects of the RI-RCLPM, for instance using time invariance for the adjacent time points t = 2, t = 3 and for t = 4, t = 5 resulting in 4 reciprocal parameters instead of 8. This makes models (a), (b), and (c) equivalent; see also the discussion in Section 2. Model (d), RI-IRCLPM, has fewer parameters and is thus not equivalent to the other three models when T = 5. Model (e) has the same number of parameters as (a), (b), (c) and is equivalent to them.

5.1.1 Model fit

The agreement in MLR model fit for the first three equivalent models of Figure 6 is demonstrated in Table 6. Note, however, that there are special considerations for the RI-RCLPM in the MWI data set. RI-RCLPM in Table 6 uses restriction (b) of non-duality and gets 2 negative R-square values which means that the solution should not be used. RI-RCLPM using restriction (a) of non-duality and positive R-square gets a worse logL = -1535 (BIC = 3357) which means that the equivalence with RI-CLPM is lost. The solution is, however, acceptable. RI-RCLPM with fully time invariant

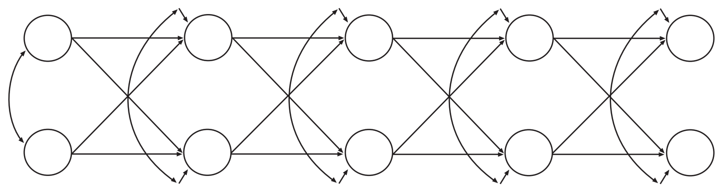
Figure 4: Mplus input for the random intercept reciprocal cross-lagged model (RI-RCLPM)

```
TITLE:          Reciprocal RI-CLPM for MWI data
DATA:           FILE = mwi.dat;
VARIABLES:      NAMES = id s1-s5 d1-d5;
                USEVAR = s1-s5 d1-d5;
                MISSING = ALL (-999);

ANALYSIS:       ESTIMATOR = ML;
                ! ML for bootstrap.
                ! Use MLR for chi-2
                BOOTSTRAP = 500;
                STARTS = 20;

MODEL:          ! Random intercepts:
                is BY s1-s5@1;
                id BY d1-d5@1;
                ! Auto-regressions:
                s2^-s5^ PON s1^-s4^;
                d2^-d5^ PON d1^-d4^;
                ! Cross-lags:
                s2^-s5^ PON d1^-d4^;
                d2^-d5^ PON s1^-s4^;
                ! Reciprocals:
                s2^-s5^ PON d2^-d5^ (rsd);
                d2^-d5^ PON s2^-s5^ (rds);
                s1^ WITH d1^;
MODEL
CONSTRAINT:     ! 2 alternatives
                ! (a) R2 pos and non-duality:
                0 <rsd*rds;
                0 <1 - rsd*rds;
                ! (b) Non-duality:
                ! 0 >(rsd*rds)^2 - 1;

OUTPUT:         STDYX RESIDUAL TECH1
                CINTERVAL(BOOTSTRAP);

PLOT:           TYPE = PLOT3;
```

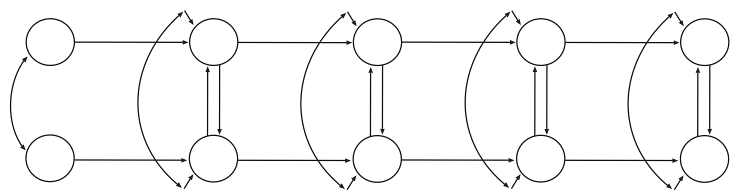
Figure 5: Mplus input for the random intercept invariant reciprocal cross-lagged model with invariant residual correlations (RI-IRCLPM)

```
MODEL:
        is BY s1-s5@1;
        id BY d1-d5@1;
        s2^-s5^ d2^-d5^ PON s1^-s4^ d1^-d4^;
        s2^-s5^ PON d1^-d4^ (sd);
        d2^-d5^ PON s1^-s4^ (ds);
        s2^-s5^ PON d2^-d5^ (r1);
        d2^-d5^ PON s2^-s5^ (r2);
        s1 WITH d1;
        ! Added statements for RI-IRCLPM:
        s2-s5 PWITH d2-d5*-.01 (c2-c5);
        s2-s5 (v2-v5); d2-d5 (w2-w5);
MODEL
CONSTRAINT:
        c3=c5*SQRT((v3*w3)/(v5*w5));
        c4=c5*SQRT((v4*w4)/(v5*w5));
        c2=c5*SQRT((v2*w2)/(v5*w5));
        NEW(t1 t2);
        r1=1/(t2*t2*c5*(1+EXP(t1)));
        r2=t2*t2*c5;
```

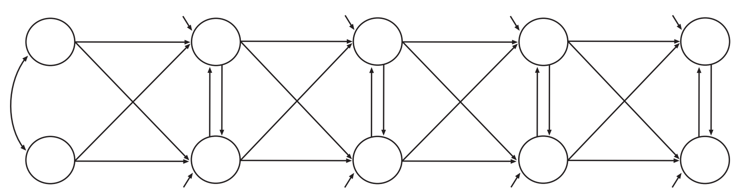
Figure 6: Five models for T = 5 (circles denote latent within-level variables)



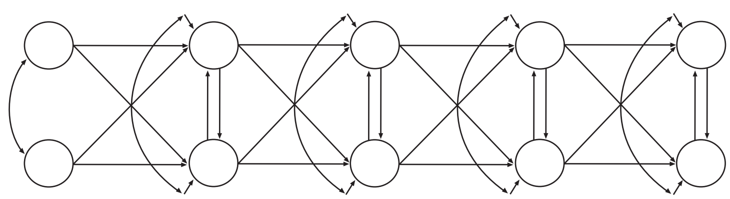
(a) RI-CLPM



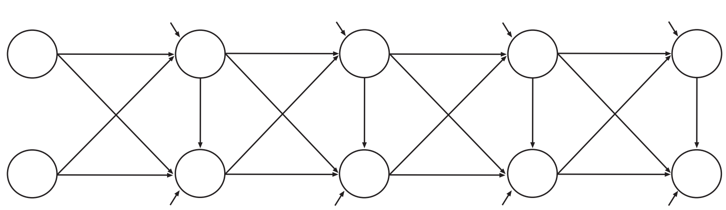
(b) RI-RPM



(c) RI-RCLPM



(d) RI-IRCLPM



(e) RI-CLPM-Lag0

Table 6: MWI model fit for three equivalent random intercept cross-lagged and reciprocal models using MLR

Model	# par's	LogL	BIC	Chi-square	Df	P-value	RMSEA	P-value
RI-CLPM	44	-1532	3349	34	21	0.0323	0.031	0.958
RI-RPM	44	-1532	3349	34	21	0.0323	0.031	0.958
RI-RCPLM 4 reciprocals: 2=3, 4=5	44	-1532	3349	34	21	0.0323	0.031	0.958

reciprocals (2 instead of 4 reciprocals estimated) gets the same logL = -1532 in these data with 2 fewer parameters and therefore a better BIC value (see Table 7).

Table 7 shows a series of nine models. Models 1 and 2 are of the RI-CLPM type, models 3 and 4 are of the RI-RPM type, models 5 - 8 are of the RI-RCLPM type, and model 9 uses the RI-IRCLPM. As stated in Section 3, it should be noted that the three model types RI-CLPM, RI-RPM, and RI-RCLPM are not nested so that chi-square difference testing is not appropriate across model type, only within model type.[6]

Model 1 is equivalent to model 3 as mentioned earlier. For these data, however, model 3 obtains negative R^2 values and is therefore disregarded. Model 4 imposes time invariance of the reciprocal effects and obtains positive R^2 values. Ignoring the adjustment for scaling correction factors in the chi-square difference testing using MLR, the results suggest that the invariance is suitable. The BIC value is the best among the first four models.

Model 1 is also equivalent to an RI-RCLPM with reciprocals restricted to equality for e.g. times 2=3, 4=5. This means that comparing model 5 to Model 1 tests full reciprocal invariance 2=3=4=5. The outcome of this test is that model 5 is not rejected because the log likelihod is the same in this data set. Imposing time invariance of the cross-lagged effects in model 2 imposes 6 restrictions (2 instead of 8 cross-lagged effects). Ignoring the adjustment for scaling correction factors in the chi-square difference testing using MLR, the results suggest that the invariance is not suitable. In contrast, imposing the time invariance restrictions on the cross-lagged effects of model 5 to obtain model 6, suggests that the invariance is suitable.[7] Model 6 also has a better BIC value than any of the previous models. Before moving to the last three models, it is of interest to look at the results of model 6.

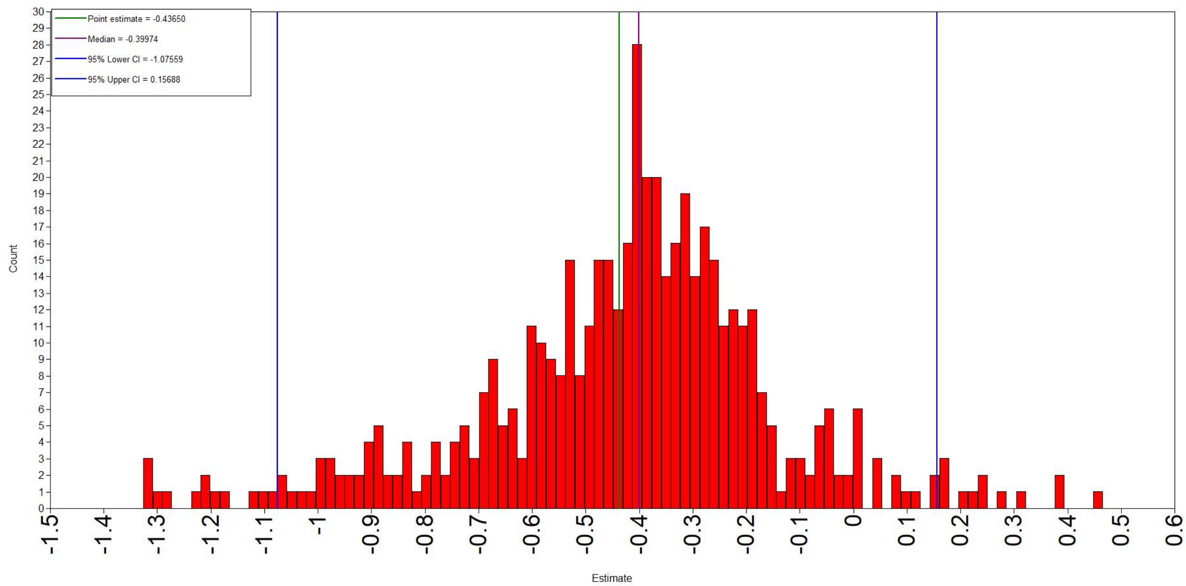
Model 6 obtains a zero estimate for the effect $\mathrm{D_t} \rightarrow \mathrm{S_t}$. In line with guideline 4 in Section 3.1.1, the recommendation is to fix this parameter at zero. This does not cause a change in the loglikelihood. Another way to arrive at this conclusion is by considering the bootstrap distributions of the reciprocal effects. Figure 7 shows the bootstrap distribution of the reciprocal effect $\mathrm{S_t} \rightarrow \mathrm{D_t}$ for the RI-RCLPM model 6.[8] The

[6]See, however, the special case of testing model 5 againt model 1 in Table 7 below.
[7]Both models 5 and 6 need restriction (a) to obtain positive R^2 values.
[8]This is obtained using the PLOT command in Mplus.

Table 7: MWI model fit for cross-lagged and reciprocal models using MLR

Model	# par's	LogL	BIC	Chi-square	Df	P-value	RMSEA	P-value
1. RI-CLPM	44	-1532	3349	34	21	0.0323	0.031	0.958
2. RI-CLPM Invar. X-lags	38	-1546	3338	60	27	0.0002	0.043	0.763
3. RI-RPM	44	-1532	3349	34	21	0.0323	0.031	0.958
4. RI-RPM Invar. Recips	38	-1538	3323	45	27	0.0181	0.031	0.975
5. RI-RCLPM Invar Recips	42	-1532	3337	34	23	0.0637	0.025	0.990
6. RI-RCLPM Invar. X-lags and Recips	36	-1539	3313	44	29	0.0409	0.027	0.992
7. RI-RCLPM Invar. X-lags and Recips $S_t \rightarrow D_t$ only	35	-1539	**3306**	45	30	0.0386	0.027	0.993
8. RI-RCLPM Invar. X-lags and Recips $D_t \rightarrow S_t$ only	35	-1546	3319	59	30	0.0011	0.038	0.906
9. RI-IRCLPM Invar. X-lags and Recips and Res. corr.	37	-1539	3319	na	na	na	na	na

Figure 7: Bootstrap Distribution of the reciprocal effect $S_t \rightarrow D_t$ for MWI model 6



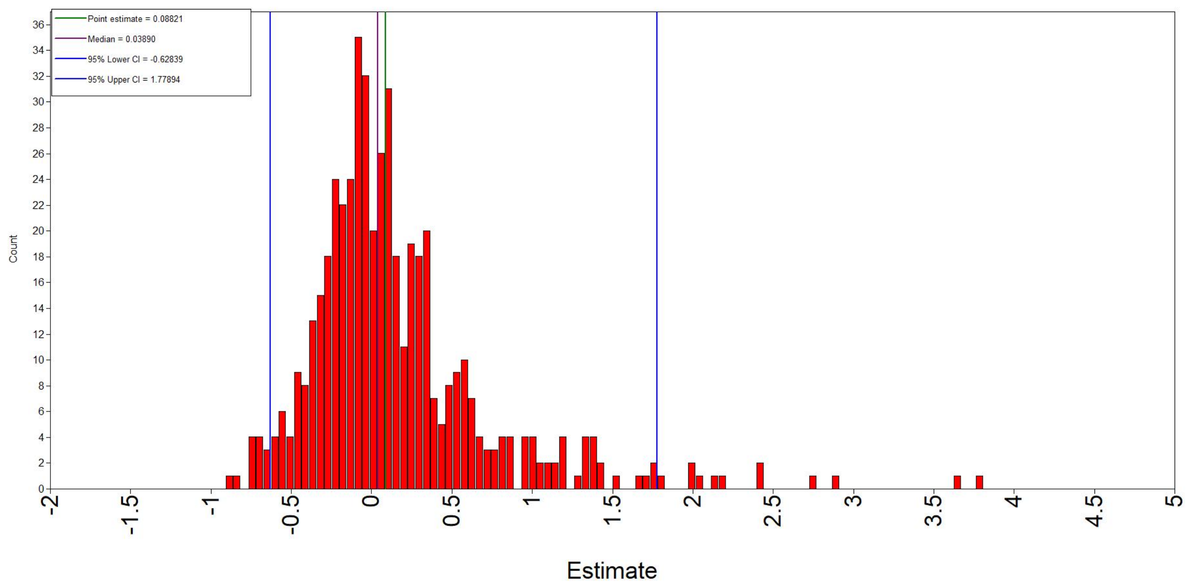
distribution has a slight skewness with a long left tail. It shows a majority of negative values with a peak around the value of -0.4. The upper limit of the confidence interval is slightly above zero so that the effect is not significant. The effect is, however, significant when eliminating the insignificant effect in the opposite direction. The $D_t \rightarrow S_t$ effect is shown in Figure 8. Here, there is a peak around zero and the parameter is not significant but can be fixed at zero.

Model 7 of Table 7 fixes to zero the contemporaneous effect $D_t \rightarrow S_t$ and only estimates the $S_t \rightarrow D_t$ effect. This is an example of the RI-CLPM-Lag0 model shown in Figure 6 (e) but the RI-RCLPM acronym is kept to emphasize where this more parsimonious model originated. Imposing this restriction is supported by model 7 obtaining the same log likelihood as model 6 for this data set. Model 7 has the best BIC among all the models in Table 7. As a check, model 8 fixes the contemporaneous effect in the other direction and allows only the $D_t \rightarrow S_t$ effect. This model gets a worse log likelihood, a worse BIC, and a worse chi-square test. Model 9 uses RI-IRCLPM which like model 7 imposes time invariance of cross-lagged and reciprocal effects but adds time-invariant residual correlations. The version that imposes the admissibility restrictions confirms the finding of model 7 in that only the $S_t \rightarrow D_t$ effect is found and the residual correlation is zero.[9] The Section 3.5 model variation 3 with only one contemporaneous effect (not shown in the table) also supports the model 7 finding with an insignificant residual correlation and with a worse fit for the contemporaneous effect in the opposite direction as was also seen in model 8.

For the MWI data, the RI-RCLPM model 7 may be preferred based on its superior BIC value. Model 7 is also more informative than the others, containing both cross-lagged and contemporaneous effects. A caveat is that the BIC advantage of model 7

[9]This run required special settings to provide a solution.

Figure 8: Bootstrap Distribution of the reciprocal effect $D_t \rightarrow S_t$ for MWI model 6



as applied to the MWI data is mostly obtained by being able to apply parsimonious versions of the RI-RCLPM in the form of full time-invariance of reciprocal effects as well as time-invariance of cross-lagged effects. An overall conclusion from the set of analyses is that there is not a large difference in model fit between the three model types. Nevertheless, the interpretation of the results from the different model types is quite different as summarized in Table 8. For the RI-CLPM model 2, there is a significant negative cross-lagged effect $D_{t-1} \overset{-}{\longrightarrow} S_t$ as was also found in Orth et al. (2021). For the RI-RPM model 4, cross-lagged effects are not included but there are significant negative contemporaneous effects of similar size in both directions which in standardized metric range from -0.3 to -0.7. For the RI-RCLPM model 7, there is a significant negative cross-lagged effect $D_{t-1} \overset{-}{\longrightarrow} S_t$ for which the standardized value of -0.1 is close to that of model 2, but there is also a significant negative contemporaneous effect in the opposite direction, $S_t \overset{-}{\longrightarrow} D_t$. It is noteworthy that while the model 7 cross-lagged effects is about -0.1 in a standardized metric, the contemporaneous effect in the opposite direction has a much larger standardized value of about -0.4, clearly leading to a different interpretation of effects than in the conventional model 2.

The MWI example demonstrates a key issue when adding models that go beyond the traditional RI-CLPM. RI-RCLPM can be used as a stepping stone to find a more parsimonious model such as model 7 that features a single-direction contemporaneous effect.

5.1.2 Indirect effects

Figure 9 shows the relationships between Self-esteem and Depression for the last two time points of the RI-RCLPM model 7. To illustrate how indirect and direct effects are formed, the effect of Self-esteem and of Depression at time 4 on Depression at time

Table 8: Estimated effects for the MWI data

Model	Significant Cross-lags	Significant Reciprocals
2. RI-CLPM Invar. X-lags	$D_{t\text{-}1} \xrightarrow{-} S_t$	NA
4. RI-RPM Invar. Recips	NA	$S_t \xrightarrow{-} D_t$ $D_t \xrightarrow{-} S_t$
7. RI-RCLPM Invar. X-lags and Recips $S_t \rightarrow D_t$ only	$D_{t\text{-}1} \xrightarrow{-} S_t$ (sig. also with bootstrap CI)	$S_t \xrightarrow{-} D_t$ (sig. also with bootstrap CI)

5 are considered.[10] The figure shows the within-level relationships where the influence of random intercepts has been controlled for. Table 9 presents the standardized effect estimates and their confidence intervals using bootstrapping. The total effect from Self-esteem at time 4 to Depression at time 5 is significant and negative. The total indirect effect of -0.339 has three components where the largest path of -0.223 goes via Self-esteem at time 5 (s4 →s5 →d5). The direct effect is insignificant (shown as a broken arrow in Figure 9). The total effect of Depression at time 4 on Depression at time 5 is significant. Apart from a small but significant indirect effect via Self-Esteem at time 5, it consists almost completely of the direct effect.

5.2 Example 2: The Ormel data on depression and disability

The data used in the Ormel et al. (2002) article is a sample of N = 753 individuals in an aging study concerning depression and disability. There are three time points with a time interval of 1 year. The depression and disability measures refer to current status. Figure 10 reproduces Figure 2 of the article and shows that their model has random intercepts and that the within-level variables have cross-lagged as well as reciprocal effects, that is, it is an example of an RI-RCLPM. A careful analysis was undertaken in the article using both forward and backward model fitting. The final model mimics that of the MWI analysis with RI-RCLPM in that one of the reciprocal effects was found insignificant.

T = 3 is the minimum number of time points for an RI-RCLPM (as well as for an RI-CLPM). As discussed earlier, with T = 3 the RI-RCLPM is equivalent to an RI-

[10] These effects are computed by MODEL INDIRECT in Mplus.

Figure 9: Indirect and direct standardized effects on Depression at time 5 using model 7 for the MWI data

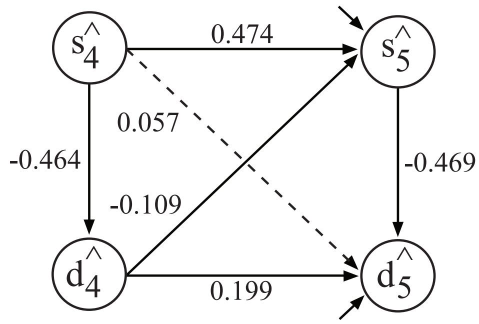
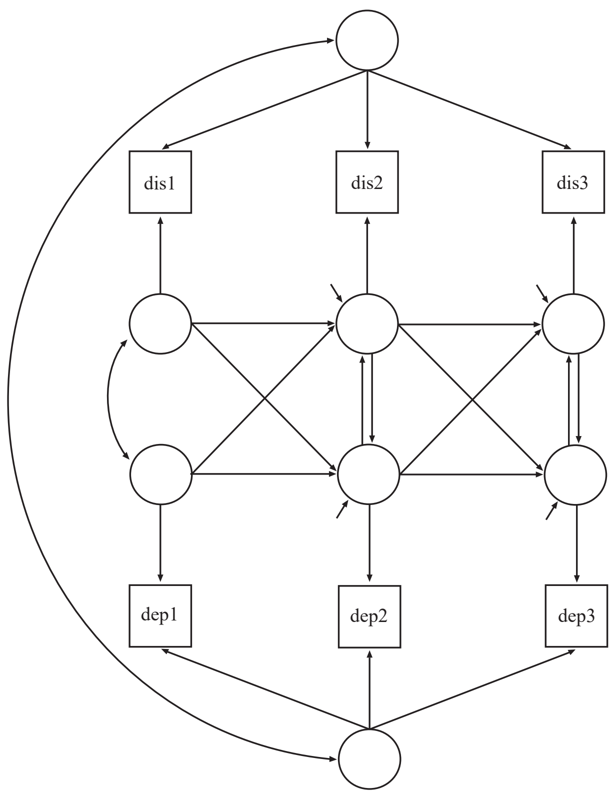


Table 9: Indirect and direct effects on Depression at time 5 using model 7

Effect	Estimate	Bootstrap Confidence Interval
Total s4 →d5	-0.282	[-0.421 - 0.148]
Total indirect s4 →d5	-0.339	[-0.459 -0.206]
s4 →s5 →d5	-0.223	[-0.302 -0.128]
s4 →d4 →d5	-0.092	[-0.157 -0.015]
s4 →d4 →s5 →d5		insignificant
Direct s4 →d5		insignificant
Total d4 →d5	0.250	[0.070 0.402]
Total indirect d4 →d5	0.051	[0.012 0.103]
d4 →s5 →d5	0.051	[0.012 0.103]
Direct d4 →d5	0.199	[0.028 0.341]

Figure 10: Model of depression and disability (Ormel et al., 2002)



CLPM when the RI-RCLPM imposes time-invariant reciprocal effects (in Ormel et al., 2002, time-invariance of both cross-lagged and reciprocal effects was imposed). The RI-RPM without time-invariant reciprocal effects is a third equivalent model. This model equivalence was not pointed out in the article but only the RI-RCLPM effects were discussed.

Raw data for this example are no longer available. Table 2 of the article, however, gives the estimated covariance matrix for the saturated model taking missing data into account using FIML (76% have complete data for all 3 time points). This estimated covariance matrix will be used as the sample covariance matrix to give parameter estimates that are likely close to what the raw data would give.[11] Chi-square test of model fit and standard errors are, however, distorted and will not be considered.

Table 10 shows seven models fitted by maximum-likelihood.[12] Model fit information is not included since it is unknown for reasons mentioned earlier. Models 1 and 2 are of the RI-CLPM type, models 3 and are of the RI-RPM type, and models 5-7 are of the RI-RCLPM type. The RI-IRCLPM with time invariant cross-lagged and reciprocal effects as well as time invariant residual correlations failed in this example as did the version with time varying residual covariances (perhaps for this reason, Ormel et al., 2002, did not include residual covariances). The Section 3.5 model variation 3 with only one contemporaneous effect did not reach convergence for either contemporaneous effect analysis.

Comparing models 1 and 2 shows that the time-invariance imposed on the cross-lagged effects in Ormel et al. (2002) was warranted. The equivalence of model 1 and model 3 is seen in the models having the same number of parameters and the same log likelihood value. Comparing models 3 and 4 shows that time invariance of the reciprocal effects fits a little worse. Comparing model 3 with model 2 shows that time invariance of reciprocals has less support than time invariance of cross-lagged effects. It should be noted that models 3 and 4 converged only with starting values derived from RI-CLPM estimates using formulas (61) - (63) of the Supplementary material.[13] These models, however, obtained negative R-square values and did not converge with restrictions (a) or (b). Hence, models 3 and 4 are not useful for this data set. Model 5 is equivalent to models 1 and 3. Model 6 imposes time-invariant cross-lagged effects which results in the same log likelihood as for model 5. Just as for RI-CLPM, the restriction of time-invariant cross-lagged effects fits the data perfectly. Model 7 eliminates the contemporaneous effect of depression on disability and finds that the log likelihood is not worsened at all, demonstrating that the eliminated effect was not needed. This results in the best BIC among the seven models. The model 7 estimates are very close to those presented in Figure 2 of the Ormel et al. (2002) article supporting the idea that analyzing the saturated estimated covariance matrix as the sample covariance matrix gives estimates that are representative of what would be obtained if the raw data were available.

Table 11 presents the estimated effects for the key models of the Ormel data set. Model 7 is the model presented in the Ormel et al. (2002) article. There is a significant

[11]This is also supported by the estimates for the final RI-RCLPM being close to those in Figure 2 of the article.

[12]MLR is not possible when analyzing a covariance matrix.

[13]The derived parameter values can also be fixed in an interim analysis to get values for the other parameters which can then be used as starting values in a subsequent run.

Table 10: Ormel data model fit (ML)

	#par's	LogL	BIC
1. RI-CLPM	20	-9139	18410
2. RI-CLPM Invar X-lags	18	-9139	18397
3. RI-RPM	20	-9139	18410
4. RI-RPM Invar Recip's	18	-9142	18403
5. RI-RCLPM Invar Recip's	20	-9139	18410
6. RI-RCLPM Invar X-lags and Recip's	18	-9139	18397
7. RI-RCLPM Invar X-lags and Recip's $\mathrm{DIS_t} \rightarrow \mathrm{DEP_t}$ only	17	-9139	**18391**

Table 11: Estimated effects for the Ormel data

Model	Significant Cross-lags	Significant Reciprocals
2. RI-CLPM Invar. X-lags	$\mathrm{DIS_{t\text{-}1}} \xrightarrow{+} \mathrm{DEP_t}$ $\mathrm{DEP_{t\text{-}1}} \xrightarrow{+} \mathrm{DIS_t}$	NA
4. RI-RPM Invar. Recips	NA	No solution
7. RI-RCLPM Invar. X-lags and Recips	$\mathrm{DEP_{t\text{-}1}} \xrightarrow{+} \mathrm{DIS_t}$	$\mathrm{DIS_t} \xrightarrow{+} \mathrm{DEP_t}$

positive cross-lagged effect from depression to disability, $\mathrm{DEP_{t\text{-}1}} \rightarrow \mathrm{DIS_t}$, and a significant positive contemporaneous effect from disability to depression, $\mathrm{DIS_t} \rightarrow \mathrm{DEP_t}$. The equivalent RI-CLPM model 2, however, concludes that both effects are cross-lagged. In this way, the two models agree about the cross-lagged effect $\mathrm{DEP_{t\text{-}1}} \rightarrow \mathrm{DIS_t}$ but disagree about the time lag for the effect from disability to depression.

5.3 Example 3: The BLS data on self-esteem and depression

The BLS data set is from the Orth et al. (2021) article which as for the MWI data concerns the relationship between self-esteem and depression. The sample size is N = 404 and T = 4 with measurements at a 1-year interval.

Table 12 presents the fit for a series of six models, with two variations for each of the three model types.[14] With T = 4, the degrees of freedom is considerably lower than for the T = 5 example of the MWI data (for T = 3 the degrees of freedom is even lower). This implies that the model is less strong in the sense of imposing fewer restrictions on the covariance matrix and is more likely to fit well as is seen in the high p-values for the chi-square tests. For RI-CLPM, time-invariance of cross-lagged effects appears supported given the small difference in log likelihood values (a formal chi-square difference test needs to take the scaling correction factors into account). For the RI-RLPM, model 3 is equivalent to model 1. Time-invariance of the reciprocal effects cannot be rejected. For RI-RCLPM, time-invariance of the cross-lagged effects

[14]The RI-IRCLPM with time invariant cross-lagged and reciprocal effects as well as time invariant residual correlations failed in this example as did the version with time varying residual covariances. The Section 3.5 model variation 3 with only one contemporaneous effect failed for both the two contemporaneous effect analyses.

Table 12: Model fit for the BLS data (MLR)

	#par's	LogL	BIC	χ^2	Df	P-value	RMSEA	P-value
1. RI-CLPM	35	-1579	3368	6	9	0.6910	0.000	0.973
2. RI-CLPM Invar X-lags	31	-1581	3349	10	13	0.6567	0.000	0.984
3. RI-RPM	35	-1579	3368	6	9	0.6911	0.000	0.973
4. RI-RPM Invar Recips	31	-1580	3347	9	13	0.7908	0.000	0.994
5. RI-RCLPM Invar Recips	34	-1581	3366	12	10	0.3048	0.021	0.871
6. RI-RCLPM Invar X-lags and Recips	30	-1582	**3344**	11	14	0.6793	0.000	0.990

cannot be rejected and this results in model 6 having the best BIC.[15]

The major impression for this example is the closeness of the log likelihood values across the models, making it hard to choose between models. The effect interpretation, however, is quite different as seen in Table 13. The RI-CLPM model 2 finds no significant cross-lagged effects. The RI-RPM model 4 finds no significant reciprocal effects. The RI-RCLPM finds a significant contemporaneous effect from depression to self-esteem. In this way, only RI-RCLPM model 6 finds any significant effects. Model 6 may be preferred over models 2 and 4 because it finds a relationship between the two variables and does not fit worse than alternative models.[16]

5.4 Example 4: The NLSY data on self-esteem and depression

The NLSY data set is also from the Orth et al. (2021) article concerning self-esteem and depression among adolecscents and young adults. Here, the sample size is much larger than for the other data sets, N = 8,259. The number of time points is also much larger, T = 11. The time interval between measurements is 2 years. The data set is characterized by having a maximum of 8 time points for any one person and has low coverage and zero coverage for several adjacent time points. This makes it impossible

[15]Models 5 and 6 results use restriction (a), needed to get positive R-square values.

[16]Unlike the MWI example, fixing $S_t \rightarrow D_t$ at its Model 6 estimate of zero, does not give a significant cross-lagged effect $S_{t-1} \rightarrow D_t$.

Table 13: Estimated effects for the BLS data

Model	Significant Cross-lags	Significant Reciprocals
2. RI-CLPM Invar X-lags	None	NA
4. RI-RPM Invar Recips	NA	None
6. RI-RCLPM Invar X-lags and Recips	None	$\mathrm{D_t} \xrightarrow{-} \mathrm{S_t}$ (sig. also with bootstrap CI)

to compute the MLR H1 model so that an MLR chi-square test is not available. ML results are instead reported. This will inflate the chi-square and underestimate the standard errors given that the outcomes are quite skewed.

Table 14 presents the fit for a series of five models, one for each of the three model types with two variations of RI-RCLPM, plus the Section 3.3 model RI-IRCLPM.[17] With T = 11, the degrees of freedom is much larger than for previous examples. Together with the much larger sample size, this tends to produce model rejection using chi-square. Due to the zero coverage for several adjacent time points, convergence is not obtained by the RI-CLPM version that has time-varying cross-lagged effects. Instead, model 1 imposed time invariance of these effects. Note that model 1 and model 2 are not equivalent despite having the same number of parameters. Equivalence holds if neither the cross-lagged nor the reciprocal effects are time invariant. As is seen in the better log likelihood value for model 2, invariance of reciprocal effects fits the data better than invariance of cross-lagged effects. For RI-RCLPM model 3, time invariance of both cross-lagged and reciprocal effects gives a somewhat worse log likelihood. There are, however, 8 fewer parameters (2 reciprocal effects instead of 10 residual covariances) so that model 3 has better BIC. As in previous examples, one of the reciprocal effects of the RI-RCLPM is insignificant and is fixed to zero in model 4. The log likelihood worsens only slightly. Because of the large sample size giving a strong penalty for using more parameters, this results in the best BIC among the models. Model 5 is the time-invariant reciprocal cross-lagged model RI-IRCLPM that includes residual covariances. Here, time invariance was imposed also on the residual correlations.[18] For the reciprocal effects, only $\mathrm{D_t} \rightarrow \mathrm{S_t}$ was significant. The residual

[17]The Section 3.5 model variation 3 with only one contemporaneous effect did not reach convergence for both contemporaneous effect analyses.

[18]The analysis had difficulty converging without making the convergence criterion somewhat more relaxed.

Table 14: Model fit for the NLSY data (ML)

Model	# par's	LogL	BIC	χ^2	Df	P-value	RMSEA	P-value
1. RI-CLPM Invar Xlags	80	-37461	75644	403	179	0.000	0.012	1.000
2. RI-RPM Invar Recips	80	-37445	75612	371	179	0.000	0.011	1.000
3. RI-RCLPM Invar Xlags Invar Recips	72	-37465	75580	412	187	0.000	0.012	1.000
4. RI-RCLPM Invar Xlags Invar Recips $D_t \rightarrow S_t$ only	71	-37466	**75572**	412	188	0.000	0.012	1.000
5. RI-IRCLPM Invar Xlags Invar Recips Invar Res corrs	73	-37466	75591	413	186	0.000	0.012	1.000

correlation was not significant.[19]

As for the previous examples, the log likelihood values are not dramatically different across the models, making it hard to choose between models. The effect interpretation, however, is again quite different as seen in Table 15. The RI-CLPM model 1 finds significant cross-lagged effects in both directions. The RI-RLPM model 2 contradicts model 1 and instead interprets these effects as reciprocal. The RI-RCLPM model 4 retains the cross-lagged effect $S_{t\text{-}1} \xrightarrow{-} D_t$ of model 1 but changes the effect of depression on self-esteem to a contemporaneous effect, $D_t \xrightarrow{-} S_t$. Model 5 which adds invariant residual correlations agrees with model 4. Model 5 provides a good check of the simplifying assumption of zero residual covariances of model 4, RI-RCLPM.

[19]The reciprocal effects and the residual covariances were all negative, fulfilling the necessary condition for an admissible solution discussed in Section 4 of the Supplementary material.

Table 15: Estimated Effects for the NLSY Data

Model	Significant Cross-lags	Significant Reciprocals
1. RI-CLPM Invar X-lags	$S_{t\text{-}1} \xrightarrow{-} D_t$ $D_{t\text{-}1} \xrightarrow{-} S_t$	NA
2. RI-RPM Invar Recips	NA	$S_t \xrightarrow{-} D_t$ $D_t \xrightarrow{-} S_t$
4. RI-RCLPM Invar X-lags Invar Recips $D_t \rightarrow S_t$ only	$S_{t\text{-}1} \xrightarrow{-} D_t$	$D_t \xrightarrow{-} S_t$
5. RI-RCLPM Invar X-lags Invar Recips Invar Res corrs	$S_{t\text{-}1} \xrightarrow{-} D_t$	$D_t \xrightarrow{-} S_t$

5.5 Example 5: The GPA data on achievement and academic self-concept

The GPA data set is from Núñez-Regueiro et al. (2021) which studied the relationship between academic self-concept and achievement of French high school students. The sample size is N = 944 and there are T = 5 time points with measurements over 6 trimesters during first and second years of high school. This results in a time interval of about 3 months except for the last 2 time points due to missing the 5th trimester.

Table 16 shows the fit of 6 models, 2 RI-CLPM, 1 RI-RPM, and 3 RI-RCLPM.[20] The RI-CLPM model 2 imposes invariance on the first 3 cross-lagged effects which have the same time distance between measurements. This appears to fit well. The 44-parameter version of RI-RPM which is equivalent to RI-CLPM model 1, obtains negative R^2 values. The model 3 version of RI-RPM that has time-invariant reciprocal effects avoids negative R^2 values by applying restriction (a).[21] BIC is better for model 3 than the first two models. For the RI-RCLPM, there is also no solution for the 44-parameter equivalent version that estimates 4 reciprocal effects, that is, applies the stipulated T-1 = 4 restrictions on the 8 reciprocal effects. Model 4 shows the RI-RCLPM version with full time invariance of reciprocals, that is, estimating only 2 reciprocal effects. Model 5 adds invariance of the first 3 cross-lagged effects in line with model 2. This changes the log likelihood very little and therefore improves BIC. Before moving to the last model, it is of interest to look at the reciprocal estimates for model 5.

Model 5 obtains a zero estimate for the contemporaneous effect Academic Self-Concept$_t$ $\rightarrow$GPA$_t$. In line with guideline 4 in Section 3.1.1, the recommendation is to fix this parameter at zero. This is also supported by the bootstrap distributions. Figure 11 shows the bootstrap distribution of the contemporaneous effect GPA$_t$ $\rightarrow$Academic Self-Concept$_t$ for model 5.[22] Most of the mass of the distribution is for positive values with a peak around +0.4 but the vertical lines of the 95 % confidence interval contain zero so that the effect is insignificant. Figure 12 shows the bootstrap distribution of the reverse contemporaneous effect Academic Self-Concept$_t$ $\rightarrow$GPA$_t$. This distribution has a peak around zero and the effect is insignificant. This effect is fixed at zero for RI-RCLPM model 6 of Table 16, obtaining a significant effect of GPA$_t$ $\rightarrow$Academic Self-Concept$_t$ as well as the best BIC of the 6 models.

As for previous examples, Table 16 shows that there is almost no difference in the log likelihood values for the 6 models so that the models cannot really be told apart. Once again, however, the interpretations of the effects are quite different as seen in Table 17 where model 2 and model 6 disagree about the lag of the effect.

[20] The RI-IRCLPM with time invariant cross-lagged and reciprocal effects as well as time invariant residual correlations failed in this example as did the version with time varying residual covariances. The Section 3.5 model variation 3 with only one contemporaneous effect failed to give admissible results for both contemporaneous effect analyses.

[21] This analysis needs a switch from the MLR estimator to the MLF estimator to avoid stoppage due to problems estimating the standard errors.

[22] Only the non-duality restriction (b) is applied here.

Table 16: Model Fit for the GPA data (MLR)

Model	# par's	LogL	BIC	χ^2	Df	P-value	RMSEA (p-value)
1. RI-CLPM Non-inv Xlags	44	-12876	26054	71	21	0.000	0.051 (.446)
2. RI-CLPM First 3 Xlags Inv	40	-12878	26029	73	25	0.000	0.045 (0.722)
3. RI-RPM Invar Recips	38	-12881	26022	93	27	0.000	0.051 (.420)
4. RI-RCLPM Non-inv Xlags Invar Recips	42	-12877	26040	71	23	0.000	0.047 (.608)
5. RI-RCLPM First 3 Xlags Inv Invar Recips	38	-12878	26016	73	27	0.000	0.043 (.836)
6. RI-RCLPM First 3 Xlags Inv Invar. Recips $GPA_t \rightarrow ASC_t$ only	37	-12878	**26009**	73	28	0.000	0.041 (.880)

Figure 11: Bootstrap distribution of reciprocal effects for model 5: $\text{GPA}_t \rightarrow$Academic Self-Concept$_t$

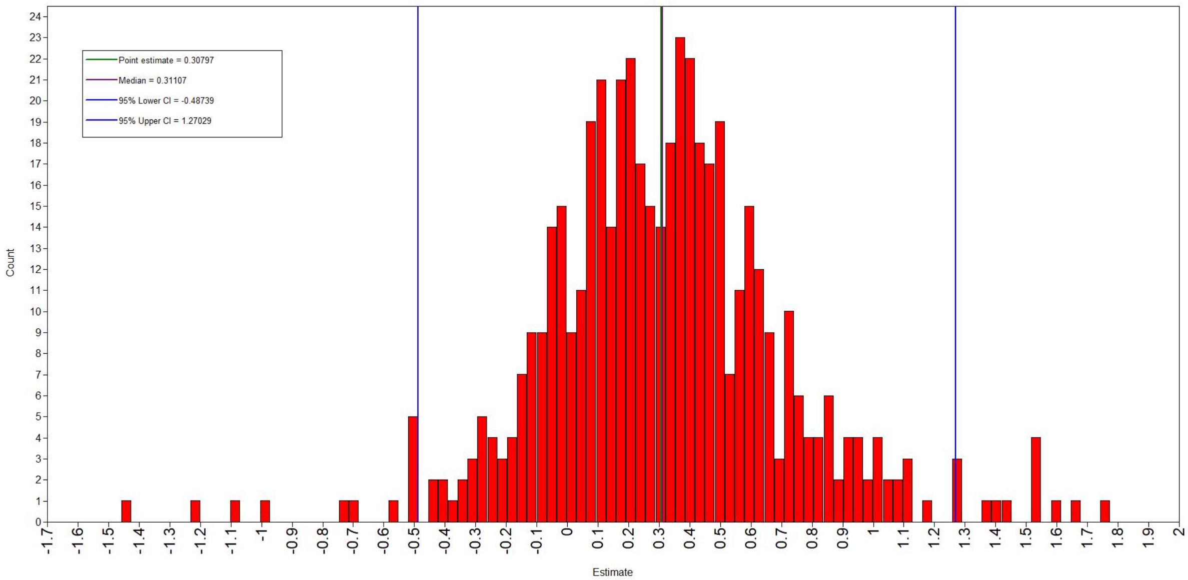


Figure 12: Bootstrap distribution of reciprocal effects for model 5: Academic Self-Concept$_t$ $\rightarrow\text{GPA}_t$

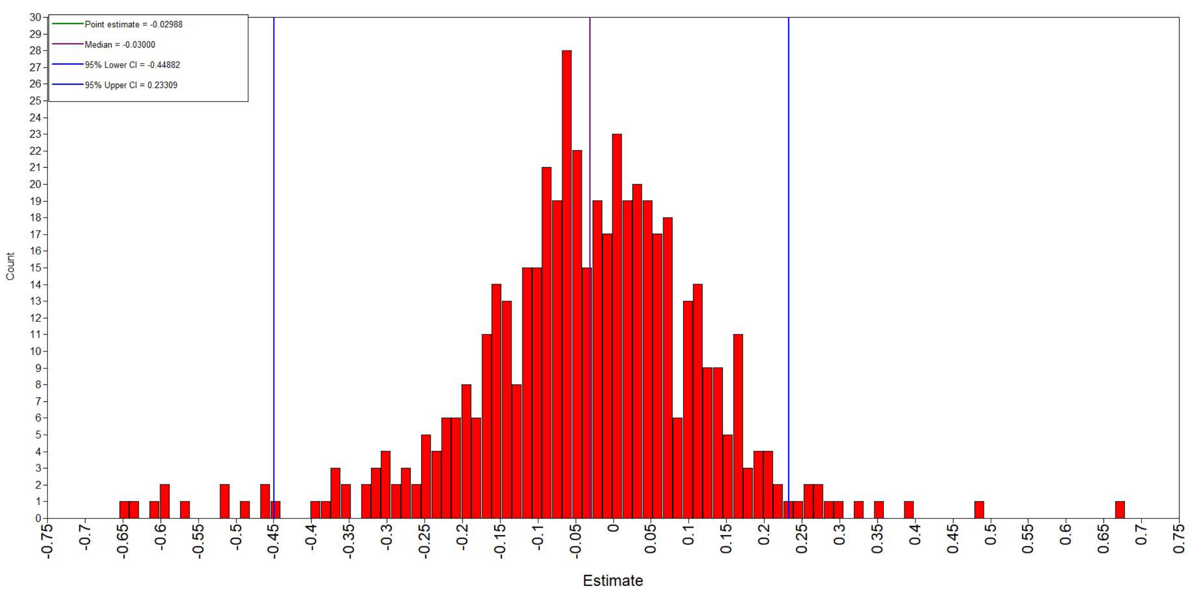


Table 17: Estimated Effects for the GPA Data

Model	Significant Cross-lags	Significant Reciprocals
2. RI-CLPM First 3 X-lags inv	$\mathrm{GPA}_{\text{t-1}} \xrightarrow{+} \mathrm{ASC_t}$	NA
3. RI-RPM	NA	$\mathrm{GPA_t} \xrightarrow{+} \mathrm{ASC_t}$
6. RI-RCLPM First 3 X-lags Inv Invar Recips $\mathrm{GPA_t} \rightarrow \mathrm{ASC_t}$ only	None	$\mathrm{GPA_t} \xrightarrow{+} \mathrm{ASC_t}$ (sig. also with bootstrap CI)

5.6 Summary of analyses

Table 18 summarizes the analysis results for the 5 examples. For the MWI data, RI-RPM challenges the finding of the regular RI-CLPM of a cross-lagged lag1 effect from depression to self-esteem and instead finds the new effects of contemporaneous (lag0) influence in both directions. RI-RCLPM supports the finding of the cross-lagged effect of the regular RI-CLPM and finds a contemporaneous effect of self-esteem on depression. For the Ormel data, RI-RCLPM supports one of the cross-lagged effects found by RI-CLPM but changes the other cross-lagged effect into a contemporaneous effect. For the BLS data, RI-RCLPM finds a contemporaneous effect whereas RI-CLPM found no effects. For the NLSY data, RI-RPM changes the RI-CLPM finding of two cross-lagged effects to two contemporaneous effects. RI-RCLPM splits the difference by changing one of the two cross-lagged effects found with RI-CLPM into a contemporaneous effect, only supporting the RI-CLPM finding of a cross-lagged effect from self-esteem to depression. For the GPA data, RI-RPM and RI-RCLPM agree in that they do not support the RI-CLPM finding of a cross-lagged effect from GPA to academic self-concept but change that into a contemporaneous effect.

Table 18 shows that for none of the 5 examples does an effect found by RI-CLPM not get found also by RI-RCLPM, albeit changing the effect from lag1 to lag0. The table also shows that using RI-RCLPM, significant reciprocal effects in both directions is not found for any of the 5 examples and for only two examples (MWI and NLSY) using RI-RPM.[23] For RI-RCLPM, a significant contemporaneous effect in one direction is found for all 5 examples.

Table 18 also provides the important finding that in the RI-RCLPM analysis of all 5 examples, a significant contemporaneous effect of $\mathrm{Y_t} \rightarrow \mathrm{Z_t}$ is accompanied by

[23] Monte Carlo simulations with substantial reciprocal effects in both directions show that such estimates are well recovered as indicated in Section 4.1.1.

a non-significant cross-lagged effect $Y_{t-1} \rightarrow Z_t$. This means that in these data sets, RI-RCLPM concludes that to predict the current status of Z, the current status of Y is a stronger predictor than the past status of Y. This is a key piece of information that the regular RI-CLPM cannot provide.

In some instances, finding a contemporaneous effect may be due to how the measurement instrument is constructed. If outcome Y_t refers to a past period and outcome Z_t refers to the present status, $Y_t \rightarrow Z_t$ can be interpreted as a lagged effect. It appears that none of the 5 examples is of this type. Perhaps the closest is the MWI example where the depression measurement refers to the past 30 days but the time frame for the self-esteem measurement is not stated. In this example, however, the contemporaneous effect is from self-esteem to depression.

6 Conclusions

This paper discusses several models that allow contemporary and reciprocal effects that have almost never been used in panel data analysis to date. The treatment of model identification, estimation, and testing shows that the reciprocal models RI-RPM and RI-RCLPM are competitors to regular RI-CLPM. Both models work well in Monte Carlo simulations. RI-RLPM worked well in 3 of the 5 examples, whereas RI-RCLPM worked well in all 5 examples. The RI-IRCLPM with time invariance for cross-lagged effects and reciprocal effects together with non-invariant residual covariances performed well in Monte Carlo simulations but proved to not work well in practice in that it failed in 3 of the 5 examples and needed special care in the other 2. The Section 3.5 model variation 3 with only one contemporaneous effect together with residual covariance also proved to not work well in practice on the 5 examples. Judging from these last two models, it appears that models that try to tease out both contemporaneous effects and residual covariances can encounter problems for the types of data with the sizes of N and T considered here. All in all, RI-RPM and RI-RCLPM appear to be the best practical alternatives to regular RI-CLPM. However, parameter restrictions on reciprocal effects are often needed to obtain admissible solutions and this makes the application of the models less straightforward. In these 5 examples, RI-RCLPM finds a significant reciprocal effect in only one direction. As illustrated in Section 4.3, this implies that RI-RCLPM can be used to decide on the direction of a contemporaneous effect in a single-direction lag0 model which is otherwise not possible given model equivalence.

As shown by the analyses of the 5 examples in Section 5, the three key model alternatives RI-CLPM, RI-RPM, and RI-RCLPM give different conclusions about the relationship between the two outcomes. The different conclusions are due to the different assumptions behind the three model types: RI-CLPM allows cross-lagged effects but assumes zero contemporaneous effects; RI-RPM assumes zero cross-lagged effects but allows contemporaneous effects; RI-RCLPM allows both cross-lagged and contemporaneous effects but in contrast to the other two model types assumes zero residual covariances. Overall for these 5 examples, the model fit is quite similar for the three model types. For some model variations there is exact model equivalence as discussed in Section 2. Therefore the statistical analysis gives little or no guidance for which set of assumptions is more reasonable. There is clearly a lack of power to distinguish between the model types.

Table 18: Summary of analyzing the 5 data sets

Model	Cross-lags	Reciprocals
MWI: N = 663, T =5, Interval = 2 months		
RI-CLPM	$D_{t\text{-}1} \xrightarrow{-} S_t$	NA
RI-RPM	NA	$S_t \xrightarrow{-} D_t$ $D_t \xrightarrow{-} S_t$
RI-RCLPM	$D_{t\text{-}1} \xrightarrow{-} S_t$	$S_t \xrightarrow{-} D_t$
Ormel: N = 753, T = 3, Interval = 1 year		
RI-CLPM	$DIS_{t\text{-}1} \xrightarrow{+} DEP_t$ $DEP_{t\text{-}1} \xrightarrow{+} DIS_t$	NA
RI-RPM	NA	No Solution
RI-RCLPM	$DEP_{t\text{-}1} \xrightarrow{+} DIS_t$	$DIS_t \xrightarrow{+} DEP_t$
BLS: N = 404, T = 4, Interval = 1 year		
RI-CLPM	None	NA
RI-RPM	NA	None
RI-RCLPM	None	$D_t \xrightarrow{-} S_t$
NLSY: N = 8,259, T = 11, Interval = 2 years		
RI-CLPM	$S_{t\text{-}1} \xrightarrow{-} D_t$ $D_{t\text{-}1} \xrightarrow{-} S_t$	NA
RI-RPM	NA	$S_t \xrightarrow{-} D_t$ $D_t \xrightarrow{-} S_t$
RI-RCLPM	$S_{t\text{-}1} \xrightarrow{-} D_t$	$D_t \xrightarrow{-} S_t$
GPA: N = 933, T = 5, Interval = 3 months		
RI-CLPM	$GPA_{t\text{-}1} \xrightarrow{+} ASC_t$	NA
RI-RPM	NA	$GPA_t \xrightarrow{+} ASC_t$
RI-RCLPM	None	$GPA_t \xrightarrow{+} ASC_t$

The overall conclusion is that there is simply not enough information in data of the type considered here to distinguish between RI-CLPM, RI-RPM, and RI-RCLPM. Because of this, analysts cannot rely on RI-CLPM to establish cross-lagged effects, nor can an analyst rely on RI-RPM or RI-RCLPM to establish contemporaneous effects. Cross-lagged effects may be seen as providing a more informative "causal" interpretation than contemporaneous effects given the time lag between cause and effect. It is therefore tempting to stay with the regular RI-CLPM. But can one really claim that cross-lagged effects have been established if a model that allows both cross-lagged and contemporaneous (lag0) effects fits the data as well and changes the lagged effects? Because of this, our answer is no to the question in the title of the paper: Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The regular RI-CLPM assumes zero contemporaneous effects whereas these effects are quite possibly present. It may be better to report results from all three model types as well as the single-direction lag0 model. The RI-RPM and RI-RCLPM are useful complements to regular RI-CLPM, enriching the understanding of the data and challenging the RI-CLPM interpretation. RI-RPM and RI-RCLPM may also facilitate the search for parsimonious models as evidenced by such models having better BIC values than RI-CLPM for the 5 examples studied here. The RI-RCLPM analysis often serves as a useful stepping stone that suggests a simple single-direction lag0 model.

The use of contemporaneous effects, single- or bi-directional, may be criticized as violating the idea of a time lag needed between cause and effect. There may, however, truly be a distinct time lag but one that is much shorter than that of the interval between measurements so that the contemporaneous model is an approximation to a model with lag somewhat greater than zero. For instance, in example 5, the two outcomes of GPA and academic self-concept refer to the same trimester, but GPA may be known before responding to the academic self-concept question. The GPA you had the previous trimester may be less relevant so that it is plausible that your current GPA most strongly influences your current academic self-concept. This was also the conclusion of the RI-RPM and RI-RCLPM analyses.

In the cross-lagged modeling overview article by Orth et al. (2021), the time intervals for the ten data sets considered were 2 months, 6 months, 1 year, and 2 years (Table 2, p. 1021). The question may be raised whether it is realistic to expect cross-lagged effects over such long time intervals or if it is more realistic that at least some of the effects between the variables are contemporaneous or approximately contemporaneous.

The analyses of this paper indicate that one can often find support for the direction of effects but due to the design of the data collection it may not be possible to determine if the lag is 1 versus 0. This was also found in Muthén and Asparouhov (2023) for an example where one of the outcomes was categorical. Perhaps a measurement design with much shorter time intervals is needed to better establish cross-lagged effects such as using intensive longitudinal data (see, e.g., Hamaker & Wichers, 2017) calling for dynamic structural equation modeling (Asparouhov et al., 2018). The choice of time interval relates to the topics of effect sizes changing as a function of time interval as discussed in Dormann and Griffin (2015) and continuous-time modeling of intensive longitudinal data (e.g., Voelkle et al., 2012; Deboeck and Preacher, 2015; Asparouhov & Muthén, 2023).

References

[1] Asparouhov, T., Hamaker, E.L. & Muthén, B. (2018). Dynamic structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, 25:3, 359-388.

[2] Asparouhov, T. & Muthén, B. (2022). Residual structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, DOI: 10.1080/10705511.2022.2074422

[3] Asparouhov, T. & Muthén, B. (2023). Continuous-time dynamic modeling in Mplus. Manuscript in preparation.

[4] Bollen, K. A. (1989). Structural Equations with Latent Variables. Oxford, England: Wiley and Sons.

[5] Deboeck, P. R. & Preacher, K. J. (2015). No need to be discrete: A method for continuous time mediation analysis. Structural Equation Modeling, 23:1, 61-75.

[6] Dormann, C. & Griffin, M. A. (2015). Optimal time lags in panel studies. Psychological Methods, 20, 489-505.

[7] Greenberg, D. F. & Kessler, R. C. (1979). A panel model of crime rates and arrest rates. American Sociological Review, vol 44, 843-850.

[8] Greenberg, D. F. & Kessler, R. C. (1982). Equilibrium and identification in linear panel models. Sociological Methods & Research, 10, 435-451.

[9] Greene, W. H. (1951). Econometric analysis. 7th edition. Pearson. New York. ISBN 0-13-139538-6

[10] Hamaker, E.L., Asparouhov, T. & Muthén, B.O. (2022). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In R.H Hoyle (Ed.), The handbook of structural equation modeling (2nd ed.): Guilford.

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

[12] Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. Current Directions in Psychological Science, 26, 10-15.

[13] Hamaker, E. L. (2023). The within-between dispute in cross-lagged panel research and how to move forward. Psychological Methods.

[14] Hayduk, L. A. (2006). Blocked-Error-R 2: A conceptually improved definition of the proportion of explained variance in models containing loops or correlated residuals. Quality & Quantity 40, 629–649.

[15] Kessler, R. C. & Greenberg, D. F. (1981). Linear Panel Analysis: Models of Quantitative Change. New York: Academic Press.

[16] Muthén, B. & Asparouhov, T. (2023). Cross-lagged panel modeling with categorical outcomes. Under review.

[17] Núñez-Regueiro, F., Juhel, J., Bressoux, P., & Nurra, C. (2021). Identifying reciprocities in school motivation research: A review of issues and solutions associated with cros-lagged effects models. Journal of Educational Psychology. Advance online publication.

[18] Ormel, J., Rijsdijk, F. V., Sullivan, M., van Sonderen, E. & Kempen, G. I. J. M. (2002). Temporal and reciprocal relationships between IADL/ADL disability and depressive symptoms in late life. Journal of Gerontology, PSYCHOLOGICAL SCIENCES, 57B, 338-347.

[19] Orth, U., Clark, D.A., Donellan, 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.

[20] Voelkle, C., Oud, J.H.L, Davidov, E. & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating Authoritarianism and Anomia. Psychological Methods, 17, 2, 176-192.

[21] Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P.D., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., Koval P., & Diener, E. (2020). From data to causes II: Comparing approaches to panel data analysis. Organizational Research Methods, Vol. 23(4), 688-716.