University of Reading

A unified framework of longitudinal models to examine reciprocal relations

Article

Accepted Version

Usami, S., Murayama, K. and Hamaker, E. L. (2019) A unified framework of longitudinal models to examine reciprocal relations. Psychological Methods, 24 (5). pp. 637-657. ISSN 1082-989X doi: 10.1037/met0000210 Available at https://centaur.reading.ac.uk/81561/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1037/met0000210

Publisher: American Psychological Association

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

Running head: A UNIFIED FRAMEWORK OF LONGITUDINAL MODELS

A Unified Framework of Longitudinal Models to Examine Reciprocal Relations

Satoshi Usami[1], Kou Murayama[2,3] and Ellen L. Hamaker[4]

1 Graduate School of Education, University of Tokyo, Japan

2 Department of Psychology, University of Reading, UK

3 Research Institute, Kochi University of Technology, Japan

4 Department of Methodology and Statistics, Utrecht University, Netherlands

Abstract

Inferring reciprocal effects or causality between variables is a central aim of behavioral and psychological research. To address reciprocal effects, a variety of longitudinal models that include cross-lagged relations have been proposed in different contexts and disciplines. However, the relations between these cross-lagged models have not been systematically discussed in the literature. This lack of insight makes it difficult for researchers to select an appropriate model when analyzing longitudinal data, and some researchers do not even think about alternative cross-lagged models. The present research provides a unified framework that clarifies the conceptual and mathematical similarities and differences between these models. The unified framework shows that existing longitudinal models can be effectively classified based on whether the model posits unique factors and/or dynamic residuals, and what types of common factors are used to model changes. The latter is essential to understand how cross-lagged parameters are interpreted. We also present an example using empirical data to demonstrate that there is great risk of drawing different conclusions depending on the cross-lagged models used.

Keywords: cross-lagged panel model, random-intercepts cross-lagged panel model, stable trait autoregressive trait and state model, latent curve model with structured residuals, autoregressive latent trajectory model, latent change score model

A Unified Framework of Longitudinal Models to Examine Reciprocal Relations

Longitudinal data analysis has been widely used to understand patterns, individual differences and relations of change in variables. One of the primary interests of researchers in longitudinal data analysis is the inference about reciprocal effects or causality between variables: How does change in one variable affect change in another variable? To study such reciprocal effects, the cross-lagged panel model (CLPM) has been extremely popular in behavioral and psychological science research.[1] The CLPM dates back to the vector autoregressive model used in time-series analysis, and it was later incorporated into the framework of structural equation modeling (SEM; Jöreskog, 1970; Jöreskog and Sörbom, 1979).

Despite the simplicity and intuitive appeal of the CLPM, solid causal conclusions require more than merely establishing lagged relations in longitudinal data that are in essence only correlational. At a superficial level, most researchers that use models with cross-lagged relations are well aware of this, and carefully avoid the term "causal" using other terms like "reciprocal" relation instead. However, an interest in the underlying causal mechanism is often the driving force behind these studies; while this is not always explicitly mentioned, it often becomes clear from the discussion section, where authors make suggestions for potential interventions, or advise certain policy changes. Such statements based on cross-lagged relations only make sense if we assume these relations reflect causal mechanisms.

A key threat in the context of correlational research is the omitted variable problem, which means that two variables may be related because they are both affected by a third, unobserved variable. In longitudinal research, there are generally two types of omitted variables: time-invariant variables, and time-varying variables. There are diverse

alternative longitudinal models that incorporate cross-lagged relations, that can control for such unobserved variables in different ways. These include the bivariate or multivariate versions of the latent change score (LCS) model (Hamagami & McArdle, 2001; McArdle & Hamagami, 2001; McArdle, 2009), the autoregressive latent trajectory (ALT) model (Bollen & Curran, 2004, 2006; Curran & Bollen, 2001), the stable trait autoregressive trait and state (STARTS) model (Kenny & Zautra, 1995, 2001), also known as the trait-state-error (TSE) model (Kenny & Zautra, 1995), the latent curve model with structured residuals (LCM-SR model; Curran et al., 2013), and the random-intercepts CLPM (RI-CLPM; Hamaker, Kruiper, & Grasman, 2015). Although there have been several studies that discussed the relations between some of these longitudinal models (e.g. Bollen & Curran, 2004; Hamaker, 2005; Hamaker et al., 2015; Usami, Hayes, McArdle, 2015, 2016; Bainter & Howard, 2016), a comprehensive comparison of these alternative models and their mathematical and conceptual relations has not yet been conducted. Consequently, it is common practice for researchers to run a single model (typically the CLPM) and evaluate its cross-lagged relations, without considering potential alternative models. In fact, applied researchers may not even be aware that some of these alternative models to examine cross-lagged relations exist.

The central aim of this article is to provide a unified statistical framework that clarifies the mathematical and conceptual relations among diverse models that include cross-lagged relations. As we will see, these models have different foci in their formulations of cross-lagged relations, giving us different interpretations of the reciprocal effects parameters. At the same time, these models share certain statistical properties, making it possible to provide a systematic and integrated overview of how these models are related to each other. The current paper serves as a first attempt to provide such a unified framework, which will hopefully help researchers in their decision making process when they want to investigate cross-lagged relations between variables.

At the beginning of this article, we discuss several existing longitudinal models that include cross-lagged relations to infer reciprocal effects. We begin with a section on models in which modeling developmental trajectories (i.e. mean structure) is not of key interest. In the subsequent section we focus on models in which modeling developmental trajectories is an important aspect. In the third section, we discuss three key features that define longitudinal cross-lagged models. A reader who is already familiar with these models may want to skip these three sections and simply consult the overview provided in Tables 2 and 3, and Figures 1 and 2. In the fourth section we present a unified framework to relate these models to each other, and discuss the key differences and similarities between all these longitudinal models. We will first explore whether the cross-lagged parameters of these models can be interpreted as causal effects from a counterfactual point of view, which is one of the dominant causal frameworks at this time. In section five, we present an example using empirical data to show the steps and key decision points in applying cross-lagged models. In the final section practical implications and potential expansion of the unified framework are discussed.

Cross-Lagged Models that do not explicitly model developmental trajectories

The CLPM is frequently used to model panel data in social and behavioral research. Panel data consist of a relatively small number of repeated measurements (at least 2, typically less than 6) from the same cases (e.g. individuals, dyads, or families). Throughout, we are interested in the inference of reciprocal effects between variables X and Y, and we let x_{it} and y_{it} be the measurements of these variables at time point t $(1 \dots t \dots T)$ for individual i $(1 \dots i \dots N)$.

The cross-lagged panel model (CLPM)

In the CLPM, x_{it} and y_{it} can be group-mean centered using:

$$\begin{aligned} x_{it} &= \mu_{xt} + x^*_{it} \\ y_{it} &= \mu_{yt} + y^*_{it}, \end{aligned} \tag{1}$$

at each time point. Here μ_{xt} and μ_{yt} are the group means at time point t, and x^*_{it} and y^*_{it} are temporal deviation terms from these group means. Note that these means are allowed to change over time; such changes may result from underlying developmental trajectories, but this is typically not the focus when the CLPM is used, and thus the CLPM implicitly removes such group-level longitudinal trajectories from the cross-lagged relations.

By definition, the deviations have a mean of zero (i.e. $E_i(x^*_{it}) = 0$ and $E_i(y^*_{it}) = 0$). The group-mean centered parts x^*_{i1} and y^*_{i1} are modeled as exogenous variables (i.e. variances and covariance of variables at $t = 1$ are separately estimated), whereas x^*_{it} and y^*_{it} for $t \geq 2$ are modeled as:

$$\begin{aligned} x^*_{it} &= \beta_{xt} x^*_{i(t-1)} + \gamma_{xt} y^*_{i(t-1)} + d_{xit} \\ y^*_{it} &= \beta_{yt} y^*_{i(t-1)} + \gamma_{yt} x^*_{i(t-1)} + d_{yit}, \end{aligned} \tag{2}$$

where β_{xt} and β_{yt} are autoregressive parameters and γ_{xt} and γ_{yt} are cross-lagged regression parameters. Note that these parameters have a subscript t, implying that they may change during the course of the study.

The cross-lagged parameters are the key parameters for inferring reciprocal relations between the variables. This cross-lagged parameter represents a simple partial regression coefficient from the predictor (e.g. $y^*_{i(t-1)}$) to the outcome variable (e.g. x^*_{it}), after controlling for the effect of the outcome variable at the previous time point (e.g. $x^*_{i(t-1)}$). This effect is often referred to as *Granger causality* (Granger, 1969).

The residuals d_{xit} and d_{yit} are usually assumed to be normally distributed and

correlated as:

$$\begin{pmatrix} d_{xit} \\ d_{yit} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega_{xt}^2 & \\ \omega_{xyt} & \omega_{yt}^2 \end{pmatrix}\right). \tag{3}$$

Here, ω_{xt}^2 and ω_{yt}^2 are the residual variances and ω_{xyt} is the residual covariance at occasion t. As another formulation of CLPM, time-invariant autoregressive (β_x, β_y) and cross-lagged parameters (γ_x, γ_y) as well as residual variances and covariances (ω_x^2, ω_y^2, ω_{xy}) can be assumed (i.e. stationarity).

The factor CLPM

An obvious way to extend the CLPM is by separating out a unique factor from the observed variables in the model, which is known as the factor CLPM. The factor CLPM has been referred to as the autoregressive cross-lagged factor model (Usami et al., 2015), the crossed-lagged regression of factors (McArdle, 2009), and quasi-simplex model (Jöreskog, 1974). In this model the observed scores are considered to consist of latent true scores (f_{xit} and f_{yit}) and unique factors (ϵ_{xit} and ϵ_{yit}, which are sometimes referred to as measurement errors), that is,

$$\begin{aligned} x_{it} &= f_{xit} + \epsilon_{xit} \\ y_{it} &= f_{yit} + \epsilon_{yit}. \end{aligned} \tag{4}$$

These unique factors are usually assumed to be normally distributed and possibly correlated, that is,

$$\begin{pmatrix} \epsilon_{xit} \\ \epsilon_{yit} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \psi_{xt}^2 & \\ \psi_{xyt} & \psi_{yt}^2 \end{pmatrix}\right). \tag{5}$$

Here, ψ_{xt}^2 and ψ_{yt}^2 are unique variances, and ψ_{xyt} is a unique factor covariance. The latent variables f_{xit} and f_{yit} are then modeled as:

$$\begin{aligned} f_{xit} &= \mu_{xt} + f_{xit}^* \\ f_{yit} &= \mu_{yt} + f_{yit}^*, \end{aligned} \tag{6}$$

where μ_{xt} and μ_{yt} are again the group means at time point t, and f^*_{xit} and f^*_{yit} are temporal deviation terms from group means, after accounting for the unique factors. The temporal deviation terms are then used in the factor CLPM, that is, for $t \geq 2$ we have

$$\begin{aligned} f^*_{xit} &= \beta_{xt} f^*_{xi(t-1)} + \gamma_{xt} f^*_{yi(t-1)} + d_{xit} \\ f^*_{yit} &= \beta_{yt} f^*_{yi(t-1)} + \gamma_{yt} f^*_{xi(t-1)} + d_{yit}, \end{aligned} \tag{7}$$

where β_{xt} and β_{yt} are autoregressive parameters, and γ_{xt} and γ_{yt} are cross-lagged regression parameters at time point t, while the unique factors have been accounted for. Hence, in the factor CLPM, the cross-lagged relations are posited between latent true scores, rather than between observed scores.

Obviously, the CLPM can be considered as a special case of the factor CLPM: If we set unique factors ϵ_{xit} and ϵ_{yit} to zero, the factor CLPM reduces to the CLPM. This means that the CLPM is nested within the factor CLPM. The factor CLPM is identified if two or more variables have been measured at three or more time points, while the CLPM requires only two time points (in which case it is saturated). However, for identification purposes, some constraints are required such as stationarity for unique factor (co)variances (i.e. ψ^2_x, ψ^2_y and ψ_{xy}). Table 1 summarizes the required number of time points to identify the various longitudinal models which are discussed in this paper.

It may seem counterintuitive that a model in which each latent variable f is measured by only one indicator is identified: It may seem impossible to distinguish between the residuals ϵ and d. However, the important point is that d affects future scores through the lagged relations, while the ϵ's are associated with only a single measurement occasion (Schuurman, Houtveen, & Hamaker, 2015). As such ϵ and d are statistically distinguishable. We further elaborate on the difference between ϵ and d later.

The random intercept cross-lagged panel model (RI-CLPM)

The RI-CLPM (Hamaker et al., 2015) can be seen as an extension of the CLPM that includes latent variables, which are referred to as random intercepts. In the RI-CLPM, x_{it} and y_{it} are modeled as

$$\begin{aligned} x_{it} &= \mu_{xt} + I_{xi} + x^*_{it} \\ y_{it} &= \mu_{yt} + I_{yi} + y^*_{it}, \end{aligned} \tag{8}$$

where μ_{xt} and μ_{yt} are the temporal group means, while I_{xi} and I_{yi} are time-invariant stable trait factors that represent an individual′s trait-like deviations from these means. Note that, while the term "trait" often refers to personality traits in the literature of psychology, in this manuscript "stable trait" simply means an invariant component of a variable in an individual over time. These stable trait factors I_{xi} and I_{yi} can be interpreted as random intercepts, and have means of 0 and a variance-covariance matrix $\boldsymbol{V}$. The residuals x^*_{it} and y^*_{it} are temporal deviations from the expected scores of individual i at occasion t (i.e. $\mu_{xt} + I_{xi}$ and $\mu_{yt} + I_{yi}$); hence they differ from x^*_{it} and y^*_{it} in Equation 2, which represent deviations from the group means. For $t \geq 2$, the cross-lagged relations are modeled as in Equation 2. The RI-CLPM reduces to the CLPM if we set the variances of I_{xi} and I_{yi} to zero, meaning there are no stable, trait-like between-person differences. This means that the CLPM is nested within the RI-CLPM. The RI-CLPM is identified if two or more variables have been measured at three or more time points (regardless of the stationarity condition).

Because the RI-CLPM separates stable between-person differences (i.e. stable trait factors) from within-person fluctuations over time, β_{xt} and β_{yt} do not represent the stability of the rank order of individuals from one time point to the next, but rather the amount of *within-person* carry-over (Hamaker et al., 2015). Similarly, the cross-lagged relations pertain to a process that takes place at the within-person level and therefore γ_x and γ_y should be interpreted as the quantities that express the extent to which the two

variables influence each other *within* individuals.[2] Hamaker et al. (2015) argued that parameter estimates in the CLPM (and the factor CLPM as well) conflate between-person and within-person processes and therefore the model provides inaccurate estimates for within-person reciprocal effects.

Note that in the RI-CLPM the initial within-person centered scores (x^*_{i1} and y^*_{i1}) are not correlated with the stable trait factors I by definition. The reason for this is that the model is based on decomposing the observed variance into between-person/trait-like variance, and within-person/state-like variance. There is no reason to assume that an individual′s temporal deviation at the first measurement from the person′s trait score is dependent on that trait score (especially not when the measurements started at a relatively arbitrary moment in time). This point also applies to the STARTS model explained below.

The stable trait autoregressive trait state (STARTS) model

By extending the CLPM with separating out unique factor from observed variables akin to the factor CLPM, and stable trait factors akin to the RI-CLPM, we obtain the STARTS model (Kenny & Zautra, 2001), which was first presented as the trait-state-error (TSE) model (Kenny & Zautra, 1995). The STARTS model was originally motivated by the latent state-trait (LST) model (Schmitt & Steyer, 1993; Steyer, Schwenkmezger, & Auer, 1990), which primarily focuses on the decomposition of trait and state variances of a single variable in longitudinal data. Although the STARTS model was initially used to investigate univariate relations, it has been extended to multivariate models, incorporating cross-lagged parameters (e.g. Luhmann, Schimmack, & Eid, 2011; Zautra, Marbach, Raphael, Lennon, and Kenny, 1995). Here we focus on the bivariate version of this model.

In the bivariate STARTS model, x_{it} and y_{it} are decomposed into latent true scores (f_{xit} and f_{yit}) and unique factors (ϵ_{xit} and ϵ_{yit}), as shown in Equation 4 (i.e.

$x_{it} = f_{xit} + \epsilon_{xit}$ and $y_{it} = f_{yit} + \epsilon_{yit}$). Again, the unique factors are usually assumed to be normally distributed (see Equation 5). Then, the latent true scores are further decomposed into individual′s expected scores ($\mu_{xt} + I_{xi}$ and $\mu_{yt} + I_{yi}$) and temporal deviations from these (f^*_{xit} and f^*_{yit}), similar to the decomposition in the RI-CLPM defined in the Equations 8. Namely,

$$\begin{aligned} f_{xit} &= \mu_{xt} + I_{xi} + f^*_{xit} \\ f_{yit} &= \mu_{yt} + I_{yi} + f^*_{yit}, \end{aligned} \tag{9}$$

Hence, the observed scores in a STARTS model can be expressed as

$$\begin{aligned} x_{it} &= \mu_{xt} + I_{xi} + f^*_{xit} + \epsilon_{xit} \\ x_{it} &= \mu_{yt} + I_{yi} + f^*_{yit} + \epsilon_{yit}. \end{aligned} \tag{10}$$

For $t \geq 2$, the temporal deviation terms (f^*_{xit} and f^*_{yit}) are modeled as in the Equation 7.

While the STARTS model adds two highly desirable features to the CLPM (i.e. unique factors and stable between-person differences), in practice the model is susceptible to improper solutions, especially when the sample size and number of time points are small. Therefore, diverse multiple indicator versions have been proposed (see Cole et al. 2005; Luhmann, Schimmack & Eid, 2011).[3]. Thus, while the STARTS model is theoretically identified if two or more variables have been measured at four or more time points, empirically it may require many more time points (cf. Kenny & Zautra, 2001). In addition, for identification purposes, stationarity constraints on unique factor (co)variances are required, like the factor CLPM.

Conclusion

The models described in this section can account for developmental trajectories (or, mean structure) over time through the changing group means (i.e. μ_{xt} and μ_{yt}). The extensions of the CLPM can be interpreted as tackling two threats to causal conclusions: a) unique factors, which we can think of as omitted time-varying variables that behave as

correlated or uncorrelated white noise processes over time (cf. the factor CLPM and the STARTS model); and b) traits, which can be thought of as omitted time-invariant variables (which are accounted for in the RI-CLPM and the STARTS model). In other words, such extensional models are not able to account for omitted variables that vary over time and that do not themselves behave as white noise (i.e. random variables like unique factors and residuals). Note that with respect to the latter, there is an assumption that the effect of the stable omitted variable is invariant over time. However, this assumption can also be relaxed by allowing the factor loadings for the common factor representing these stable between-person differences to vary over time. This suggests that, while the predicted rank-order of individuals based on the common factor is invariant over time, the distance between individuals may increase and/or decrease over time.

Cross-Lagged Models that explicitly model developmental trajectories

A popular model for investigating individual difference in developmental trajectories is the latent curve model (LCM), which was first proposed by Meredith and Tisak (1984, 1990). In this model, time is used as a predictor and individuals can have different intercepts and difference slopes (i.e. regression coefficients), resulting in different trajectories over time. Below, we discuss two ways in which the LCM has been combined with cross-lagged relations. In addition, we discuss an alternative model that has been proposed to infer reciprocal relations in the presence of developmental trajectories.

The latent curve model with structured residuals (LCM-SR)

One way in which the LCM and the CLPM have been combined, is now referred to as the LCM-SR (Curran et al., 2013; see also Chi & Reinsel, 1989; Hamaker, 2005). In this model, the trend part is separated from the cross-lagged part, such that there is a LCM with residuals that are modeled using a CLPM.

Hence, the LCM-SR that consists of a (linear) LCM, can be expressed as

$$\begin{aligned} x_{it} &= I_{xi} + (t-1)S_{xi} + x^*_{it} \\ y_{it} &= I_{yi} + (t-1)S_{yi} + y^*_{it}, \end{aligned} \tag{11}$$

where I_{xi} and I_{yi} represent the intercepts of the regression line that describes the developmental trajectories of individual i over time in variables x and y respectively, and S_{xi} and S_{yi} represent the linear slopes of these individual regression lines. Hence, these intercept and slope factors are similar to the *growth factors* of an LCM. The residuals x^*_{it} and y^*_{it} can be thought of as detrended variables with respect to individual linear growth curves, and for $t \geq 2$ they are modeled using Equations 2. In the LCM-SR the initial within-person centered/detrended scores (x^*_{i1} and y^*_{i1}) are not correlated with growth factors (I and S). Note that both the RI-CLPM and the LCM-SR include the common factor I that represents random intercepts. However, unlike the LCM-SR, the RI-CLPM does not explicitly model developmental trajectories and it simply assumes time-variant expected scores for each individual (i.e. $\mu_{xt} + I_{xi}$ and $\mu_{yt} + I_{yi}$). This point will be addressed later. For this reason, we refer to the common factor in the RI-CLPM as a stable trait factor, and that in the LCM-SR as a growth factor, respectively, although these are expressed by the same notation I in the current paper.

The autoregressive latent trajectory (ALT) model

The ALT model was proposed by Curran and Bollen (2001) with the aim to synthesize the CLPM and the LCM (see also Bollen & Curran, 2004; 2006). The bivariate ALT can be expressed as

$$\begin{aligned} x_{it} &= A_{xi} + (t-1)B_{xi} + \beta_{xt}x_{i(t-1)} + \gamma_{xt}y_{i(t-1)} + d_{xit} \\ y_{it} &= A_{yi} + (t-1)B_{yi} + \beta_{yt}y_{i(t-1)} + \gamma_{yt}x_{i(t-1)} + d_{yit}, \end{aligned} \tag{12}$$

for $t > 1$. Diverse modeling options exist for $t = 1$ because there is a recursiveness in the ALT and then the process needs to be "started up"(Bollen & Curran; 2004; Hamaker et

al., 2015). As Bollen and Curran (2004) argued, the first observation is typically not the beginning of the process, and it is thus reasonable to assume that the first observation is also predictable from unobserved realizations of the process prior to it. This suggests that at $t = 1$ we should not only account for the direct effects of the factors A and B, but also their indirect effects through past realizations of the process. To solve this problem, Curran and Bollen (2001) propose two solutions: Either the first observation (x_{i1} and y_{i1}) is treated as exogenous and allowed to be correlated freely with factors A and B, or nonlinear constraints are imposed on the loadings for the first occasion. Jongerling and Hamaker (2011) called the former and the later approaches "predetermined" and "constrained" approaches, respectively, and the constrained approach can be thought of as a special case of a predetermined approach. The most commonly used option is "predetermined", and in this approach x_{i1} and y_{i1} are assumed to be correlated with each other and the common factors A and B. On the other hand, other cross-lagged models discussed in this paper implicitly assume that the initial observations are strictly exogenous (e.g. uncorrelated with the stable trait factors at any time point). The treatment of the initial observations is an important theoretical and practical problem for some cross-lagged models like the ALT model that include common factors but do not separate stable between-person differences from within-person fluctuations over time.

The common factors A_{xi} and A_{yi} have a constant direct effect over time, while the latent variables B_{xi} and B_{yi} have changing effects over time, due to the fixed factor loadings $(t - 1)$. This specification is similar to those for the slope factor in a (linear) LCM. This assumption can be further relaxed by freely estimating the factor loadings of B_{xi} and B_{yi} (i.e. $\alpha_{xt}B_{xi}$ and $\alpha_{yt}B_{yi}$). As such, B_{xi} and B_{yi} are factors whose effects usually increase as time progresses. Due to their indirect effects through previous realizations of the process, however, the effects of these factors accumulate over time, and they actually cannot be interpreted as the intercept and slope factors (i.e. I and S

factors) in the LCM or the LCM-SR. We shall call the factors A and B in the ALT *accumulating factors*. In the later section we elaborate on this point further. The ALT model can be identified if two or more variables are measured at five or more time points. When stationarity of parameters can be assumed, the minimum number of required time points becomes four.

Note that the difference between the growth factors and the accumulating factors dissipates when the autoregressive and cross-lagged paths are not present (i.e. there are no indirect effects accumulating from the previous time points). Indeed, the ALT model can be reduced to the LCM when β and γ are restricted to 0 (Bollen & Curran, 2004): In this particular case, the factors A_{xi} and A_{yi} are equivalent to the random intercept factors of the LCM (I factor), while B_{xi} and B_{yi} represent the random slopes (i.e. S factor). On the other hand, despite the superficial similarities, the relation between the ALT model and the CLPM is more complicated, which we will address in a later section.

Although the ALT model is more popular than the LCM-SR in applied research, Curran et al. (2013) expressed a preference for the LCM-SR over the ALT model, because it clearly separates between-person differences in growth trajectories (growth factors), from a within-person reciprocal process. However, Hamaker (2005) showed that under particular circumstances (e.g. time-invariant autoregressive and cross-lagged parameters), the LCM-SR and the ALT model are statistically equivalent; in that case, the cross-lagged parameters of these two models are identical, while the growth factors scores in the LCM-SR can be expressed as a function of autoregressive/cross-lagged parameters as well as accumulating factor scores of the ALT model. Hence, in both the LCM-SR and the ALT model lagged estimates are not contaminated by between-person differences if models are correctly specified. Rather, the advantage of the LCM-SR over the ALT model is that the growth parameters have a clearer interpretation than the accumulating parameters from the ALT model: In the LCM-SR growth parameters can be interpreted as the

intercept and slope similarly to these parameters in the LCM.

The latent change score (LCS) model

The LCS model –also known as the latent difference score model–was developed by McArdle (2001) as a very general framework that encompasses many longitudinal models, including the CLPM and the LCM. Here we focus on the LCS model as it is commonly used in the literature. In the bivariate LCS model, x_{it} and y_{it} are first decomposed into latent true scores f_{xit} and f_{yit} and unique factors ϵ_{xit} and ϵ_{yit} using Equation 4. Note that in many of the publications on the LCS model, additional latent variables are used for the first time point, which are then referred to as the intercepts (e.g. McArdle, 2009). However, these additional latent variables coincide with the latent true scores at the first time point and thus do not serve any particular role (i.e. they are phantom variables), and can be omitted from the presentation without any consequences (e.g. Usami et al., 2015).

For $t \geq 2$ the latent true scores are modeled using latent change scores Δf_{xit} and Δf_{yit}, which express increments in the latent scores between adjacent time points $t-1$ and t, that is

$$\begin{aligned} f_{xit} &= f_{xi(t-1)} + \Delta f_{xit} \\ f_{yit} &= f_{yi(t-1)} + \Delta f_{yit}. \end{aligned} \tag{13}$$

Subsequently, the latent change scores are modeled as a function of a factor A, the preceding latent scores, and a residual, that is,

$$\begin{aligned} \Delta f_{xit} &= A_{xi} + \beta_x f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + d_{xit} \\ \Delta f_{yit} &= A_{yi} + \beta_y f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + d_{yit}. \end{aligned} \tag{14}$$

The factor A_{xi} and A_{yi} are typically referred to as slope factors in the LCS model literature. However, if we plug Equation 14 into 13, we obtain the following expression

$$\begin{aligned} f_{xit} &= A_{xi} + \beta^*_x f_{xi(t-1)} + \gamma_x f_{yi(t-1)} + d_{xit} \\ f_{yit} &= A_{yi} + \beta^*_y f_{yi(t-1)} + \gamma_y f_{xi(t-1)} + d_{yit}, \end{aligned} \tag{15}$$

where the autoregressive coefficients are functions of the proportional change coefficients β_x and β_y in Equation 14, that is, $\beta_x^* = (1 + \beta_x)$ and $\beta_y^* = (1 + \beta_y)$. Therefore, despite its name, slope factors in the LCS model can actually be interpreted as the baseline or intercept in the equation, similar to the factors A in the ALT model (cf. Usami et al., 2015). The cross-lagged parameters γ_x and γ_y are often referred to as coupling parameters in the LCS model.

Although the LCS model typically assumes time-invariant autoregressive and cross-lagged regression parameters, the assumption of stationarity can be relaxed and time-varying parameters (i.e. β_{xt}, β_{yt}, γ_{xt} and γ_{yt}) can also be estimated, just as in the preceding models. However, for identification purposes, stationarity constraints on unique factor (co)variances are required, like in the factor CLPM and the STARTS model. In addition, the model assumption can be further relaxed by freely estimating the factor loadings α_{xt} and α_{yt} for the factors (i.e. $\alpha_{xt}A_{xi}$ and $\alpha_{yt}A_{yi}$) to express different amounts of baseline change at each time point or to account for unequal spacing of observed data. This model is called the triple change score (TCS) model (McArdle & Nesselroade, 2014). The residuals d_{xit} and d_{yit} are usually assumed to be normally distributed as in Equation 3, and time-invariant residual variances and covariance are typically assumed in the LCS model. Moreover, for reasons of empirical identification, these (co)variances are often fixed to zero, which suggests the LCS model is used to describe the latent and deterministic trends. Similar to the (predetermined) ALT model where the observations at the first occasion are treated as exogenous and are allowed to be correlated with the growth factors, the latent scores of the LCS model at the first occasion (f_{xi1} and f_{yi1}) are treated as exogenous and allowed to be correlated freely with the factors A.

Although the relation between the LCS model as expressed in Equations 13 and 14 and other longitudinal models may be difficult to see at first, the representation in Equation 15 helps to clarify some of the connections between them. For example, the

formulation in Equation 15 shows that the model is in fact closely related to an ALT model. Specifically, when unique factors are assumed in the ALT model, this general version of ALT model is reduced to the LCS model by setting $(t-1)B_{xi} = 0$ and $(t-1)B_{yi} = 0$ (Bollen & Curran, 2004). Thus, it becomes clear that the critical difference between the LCS model and the ALT model lies in which accumulating factors are included (i.e. the ALT model has additional accumulating factors, B_{xi} and B_{yi}).

Previous research has addressed the relation between the LCS model and the LCM and the CLPM. For example, when β and γ as well as the residual variances and covariances are restricted to 0 under time-variant variances of the unique factors, the model is mathematically equivalent to the (linear) LCM (cf. Usami et al., 2015). Here, the factors A_{xi} and A_{yi} are equivalent to the random slope factors of the LCM, whereas f_{xi1} and f_{yi1} then represent the random intercepts. Usami et al. (2015) also showed that the TCS model (a general form of the LCS model) is reduced to the factor CLPM by setting A_{xi} and A_{yi} to constant values, which is mathematically equivalent to fixing the variances of these two factors (and the five covariances associated with them) to zero.

Note that Allison (2005) also introduced a cross-lagged model that is closely related to the LCS model: If we omit the unique factors ϵ from the LCS model (while keeping the residuals d), the LCS model is mathematically equivalent to Allison′s model. Like Allison′s (2005) model, there are in fact various longitudinal models in the econometric literature that account for individual-specific effects through the inclusion of a single unobserved variable like the accumulating factor (A) (e.g. Hsiao, 2014). Such models are referred to as dynamic panel models. However, to the best of the authors′ knowledge, little attention has been paid to how dynamic panel models are related to the longitudinal models discussed here among psychometric and econometric researchers.

Conclusion

The models described in this section have been explicitly proposed as ways to study reciprocal effects in the context of developmental trajectories and individual differences therein. Two of these models control for the effect of time by including time as a predictor, and allowing it to have different effects for different individuals (random slopes: i.e. B factor in the ALT model and S factor in the LCM-SR). Moreover, two of the models allow for developmental trajectories to result from the reciprocal connections between the two variables (i.e. the ALT model and the LCS model). These models tackle threats to conclusions about causality by accounting for both omitted time-varying and time-invariant variables in different ways (i.e. unique factors in the LCS model, growth factors in the LCM-SR and accumulating factors in the ALT and the LCS models). However, if there are omitted time-varying variables characterized by autocorrelation over time, these are not properly accounted for in these models and this is likely to cause biased results. With respect to omitted time-invariant variables, our model specification focused on linear effects of time (i.e. the factor loadings for the common factors B and S were set to $(t-1)$ in the ALT model and the LCM-SR respectively). However, it is also possible to allow these factor loadings to be estimated freely, such that we can account for non-linear time-variant effects. Furthermore, these models can be extended with quadratic and cubic terms to allow for more flexibility regarding the shape of trends.

Understanding the Relation Between Different Types of Longitudinal Cross-lagged Models: A Unified Framework

Based on the earlier discussion of diverse longitudinal models with cross-lagged relations, we can now identify key features of these models that define the fundamental differences and similarities between them. These key features are useful to understand the potential of longitudinal models to assess causal effects, which we will discuss later. After

describing the features, we provide a unified statistical framework that clarifies the mathematical and conceptual relations among diverse cross-lagged models.

Key features of cross-lagged panel models

To see how the diverse cross-lagged panel models are related and how they differ, it is helpful to consider the following key features: a) whether unique factors are included or not; b) the type of common factors that are included; and c) the type of trajectories they can describe. We elaborate on each feature below.

Unique factors. A key feature of the models considered here is the type of residuals they include. The factor CLPM, the STARTS model, and the LCS model all include unique factors, which affect the observed scores only at a single occasion. In contrast, the CLPM, the RI-CLPM, the ALT model, and the LCM-SR do not include unique factors. For instance, the defining difference between the CLPM and the factor CLPM is that the latter includes unique factor; the same is true when comparing the RI-CLPM and STARTS model.

All the models considered here include residuals that affect subsequent scores through the lagged relations: These are the residual terms d_{xit} and d_{yit} in the cross-lagged regression equations, which represent the parts that cannot be explained by the lagged relations (and other common factors) in the model. Because these residuals are incorporated as a structural part of the model, their influences feed forward through the lagged relations. As such, they are commonly referred to as *innovations* or *dynamic errors*, to emphasize how their effect differs from that of unique factors, which are sometimes referred to as *measurement errors*.

Whilst the inclusion of both unique factors *and* dynamic residuals is desirable based on conceptual grounds, it can easily lead to estimation problems due to the strong dependency between the estimated parameters (in the case of the STARTS model, see

Lüdtke, Robitzsch & Wagner, 2017 for more details). For example, it has been stated that STARTS models may require 10 or more time points of data to obtain reliable estimates (Kenny & Zautra, 2001). Similarly, the inclusion of both kinds of errors in the LCS model is prone to lead to convergence problems and improper solutions (e.g. McArdle et al., 2004; Usami et al., 2015). Therefore, as noted earlier, dynamic errors are often omitted from the LCS model in practical applications of it.

Common Factors. All the models discussed above, with the exception of the CLPM and the factor CLPM, contain common factors that are used to model some form of individual differences in trait level or developmental trajectory. Just as with unique factors and dynamic residuals discussed above, the role of common factors critically depends on whether or not they are separated from the lagged relations in the model. When the factors are separated from the lagged relations, as is done in the RI-CLPM, the STARTS model, and the LCM-SR, this separation suggests that the factors only have *direct* effects on (latent) scores at each occasion. In contrast, when the factors are *not* separated from the lagged relations, as is the case in the ALT and the LCS models, their effects accumulate through the lagged relations, such that they have both *direct* and *indirect* effects on the (latent) scores. This is why we call these factors (A and B) *accumulating* factors.

To illustrate this point, let us compare the STARTS model and the LCS model: Both models include unique factor and dynamic residual, and both models contain a single common factor for each variable (i.e. I_{xi} and I_{yi} in the STARTS model, and A_{xi} and A_{yi} in the LCS model). We focus on the latent score of the variable x at the third occasion (i.e. f_{xi3}).

When considering the STARTS model, we can combine Equations 7 and 9, and use

this to obtain an expression of the latent true score at occasion 3, that is,

$$\begin{aligned} f_{xi3} &= \mu_{x3} + I_{xi} + f^*_{xi3} \\ &= \mu_{x3} + I_{xi} + \beta_{x3} f^*_{xi2} + \gamma_{x3} f^*_{yi2} + d_{xi3} \\ &= \mu_{x3} + I_{xi} + \beta_{x3}[\beta_{x2} f^*_{xi1} + \gamma_{x2} f^*_{yi1} + d_{xi2}] + \gamma_{x3}\big(\beta_{y2} f^*_{yi1} + \gamma_{y2} f^*_{xi1} + d_{yi2}\big) + d_{xi3}. \end{aligned} \tag{16}$$

This expression shows that the stable trait factor I_{xi} only has a direct effect on f_{xi3}, since it does not show up in the lagged terms in the Equation. As a result, the expected score for an individual at $t = 3$ is equal to $\mu_{x3} + I_{xi}$ (i.e. the group mean for $t = 3$ plus the individual′s stable deviation from the group mean).

In contrast, in the LCS model, it can be shown that the common factor A_{xi} has both a direct and indirect effects on f_{xi3}. Using Equation 15, we can write

$$\begin{aligned} f_{xi3} &= A_{xi} + \beta^*_x f_{xi2} + \gamma_x f_{yi2} + d_{xi3}, \\ &= A_{xi} + \beta^*_x(A_{xi} + \beta^*_x f_{xi1} + \gamma_x f_{yi1} + d_{xi2}) + \gamma_x(A_{yi} + \beta^*_y f_{yi1} + \gamma_y f_{xi1} + d_{yi2}) + d_{xi3}, \\ &= (1 + \beta^*_x)A_{xi} + \gamma_x A_{yi} + \beta^*_x\big(\beta^*_x f_{xi1} + \gamma_x f_{yi1} + d_{xi2}\big) + \gamma_x\big(\beta^*_y f_{yi1} + \gamma_y f_{xi1} + d_{yi2}\big) + d_{xi3}. \end{aligned} \tag{17}$$

This shows that the effect of the common factor A_{xi} is different at each occasion, and that it becomes increasingly more complex over time. At time point $t = 2$, because f_{xi2} can be expressed as $f_{xi2} = A_{xi} + \beta^*_x f_{xi1} + \gamma_x f_{yi1} + d_{xi2}$, the effect of the common factor appears as A_{xi}. However, as shown in Equation 17, at $t = 3$, the effect becomes $(1 + \beta^*_x)A_{xi}$, which reflects both a direct effect and an indirect effect. Furthermore, this expression also shows that the expected score for person i on variable x at time point $t = 3$ is a function of: a) the common factors A_{xi} and A_{yi}; b) the lagged parameters β^*_x and γ_x, as well as β^*_y and γ_y; and c) the initial latent states f_{xi1} and f_{yi1}.

Hence, we argue that it is helpful to differentiate between models in which the common factors are modeled separately from the lagged effects (i.e. the RI-CLPM, the

STARTS model, and the LCM-SR), versus models in which the common factors have accumulated effects over time on the (latent) scores through the lagged relations (i.e. the ALT model and the LCS model). While the interpretation of factors with only direct effects is straightforward, the interpretation of the accumulative factors with their direct and indirect effects is more challenging. This issue is also relevant to the different interpretations of the cross-lagged parameters (γ), which we will discuss in the later section.

Mean structure and trends. Another important characteristic of the models discussed here is how the mean structure is modeled. In the first category of models discussed in this paper (i.e. the CLPM, the factor CLPM, the RI-CLPM, and the STARTS model), the group means are typically not constrained over time. As a result, the average trajectory can take on any shape. In addition, in the CLPM and the factor CLPM, there are no individual trajectories modeled. In the RI-CLPM and the STARTS model, the stable trait factor I suggests that individuals are characterized by a constant deviation from this average trajectory. Hence, if on average there is a linearly increasing trend, each individual is assumed to have the same trajectory that is simply shifted up or down by a constant that is captured by the stable trait factor. Similarly, if there is on average a logistic curve over time, each individual has their own logistic curve that is parallel to this average curve.

In the second category of models considered in this paper (consisting of the LCM-SR, the ALT model, and the LCS model), the average developmental trajectories are modeled with the means of the common factors I and S, or A and B. This again suggests that the way these factors are included in the model is of vital importance. As the LCM-SR includes common (i.e. growth) factors separately from the lagged effects in the model, their means operate in the same way as in the LCM: They represent the average intercept and average (linear) slope of developmental trajectories.

In contrast, in the ALT model and the LCS model the latent means are not separated from the lagged relations in the model. As a result, the average trajectory is the result of the means of common (i.e. accumulating) factors, but also of the lagged relations (i.e. β and γ). At times, this can lead to unexpected trajectories. For example, Jongerling and Hamaker (2011) showed how an ALT model with only a constant common factor A (i.e. no slope factor B), may result in an increasing or decreasing trend (or up and down when time-varying autoregressive parameters are assumed) over time. In fact, it is exactly this entanglement of the common factors A and the lagged relations that forms the basis of modeling trajectories with the LCS model.

This highlights an interesting difference in modeling traditions: Whilst the LCM-SR seems to follow common practice in time series analysis, in which data are first detrended before one investigates whether there are lagged relations within and between variables, the ALT model and the LCS model seem more in line with the dynamic system′s literature, in which short-term fluctuations are described using the same set of equations as the ones used for long-term behavior of a system.

The latter observation may lead readers to conclude that the choice between these models should depend on whether there is one set of underlying rules that govern the system, or instead that there is a developmental process that operates separately from the short-term reciprocal effects. This is partially correct. However, it is interesting that under certain conditions (e.g. time-invariant autoregressive and cross-lagged parameters), the lagged relations of these two approaches are actually identical (for details see Hamaker, 2005).

A unified framework

Based on the key features discussed above, we can now provide a unified framework that integrates all these cross-lagged longitudinal models into a general statistical model.

Path diagrams of this framework as well as cross-lagged longitudinal models we have discussed are provided in Figure 1. The unified framework makes it easy to see how the diverse models differ from one another. The formulation consists of three sets of equations, which we refer to as the *measurement equations*, the *decomposition equations*, and the *dynamic equations*.

Measurement Equations. The first set of equations can be used to separate the latent true scores from unique factors. We rewrite these measurement equations as

$$\begin{aligned} x_{it} &= f_{xit} + \epsilon_{xit} \\ y_{it} &= f_{yit} + \epsilon_{yit}. \end{aligned} \tag{18}$$

For the factor CLPM, the STARTS model and the LCS model, these equations are used; for all the other models we discussed (i.e. the CLPM, the RI-CLPM, the ALT model, and the LCM-SR), no unique factors are included, and these equations are reduced to $x_{it} = f_{xit}$ and $y_{it} = f_{yit}$.

Decomposition Equations. The second set of equations allow for a decomposition into an individual deterministic trend and a temporal deviation from this individual trend, denoted as f^*_{xit} and f^*_{yit}. The individual deterministic trend can depend on the group means μ_{xt} and μ_{yt} and/or on the random intercepts and slopes (i.e. I_{xi}, I_{yi}, S_{xi} and S_{yi}). Thus, we have

$$\begin{aligned} f_{xit} &= [\mu_{xt} + \left\{I_{xi} + (t-1)S_{xi}\right\}] + f^*_{xit} \\ f_{yit} &= [\mu_{yt} + \left\{I_{yi} + (t-1)S_{yi}\right\}] + f^*_{yit}. \end{aligned} \tag{19}$$

In the CLPM and the factor CLPM, the decomposition equations are simply used to center the data per time point (i.e. detrend the group trajectory by including μ_{xt} and μ_{yt}), whereas the other parts for the individual deterministic trends (i.e. $\left\{I_{xi} + (t-1)S_{xi}\right\}$ and $\left\{I_{yi} + (t-1)S_{yi}\right\}$) are dropped. In the RI-CLPM and the STARTS model, these decomposition equations are used to center the data per time point *and* per individual

(i.e. the expected scores of individuals are $\mu_{xt} + I_{xi}$ and $\mu_{yt} + I_{yi}$), while the remaining elements of the individual deterministic trends (i.e. $(t-1)S_{xi}$ and $(t-1)S_{yi}$) are dropped.

In contrast, the LCM-SR aims to directly evaluate growth trajectories through growth factors rather than to detrend the overall growth by using μ_{xt} and μ_{xt}. Hence in this model μ_{xt} and μ_{yt} are set to 0, and the means are modeled using the means of growth factors I and S. Finally, the ALT model and the LCS model do not make use of the options offered by the decomposition equations, meaning $f_{xit} = f^*_{xit}$ and $f_{yit} = f^*_{yit}$.

Dynamic Equations. Finally, the dynamics of the processes are modeled with the *dynamic equations*, which include the lagged relations over time. In addition, they also include the accumulating factors A and B. This gives

$$\begin{aligned} f^*_{xit} &= \left\{A_{xi} + (t-1)B_{xi}\right\} + \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit} \\ f^*_{yit} &= \left\{A_{yi} + (t-1)B_{yi}\right\} + \beta_y f^*_{yi(t-1)} + \gamma_y f^*_{xi(t-1)} + d_{yit}. \end{aligned} \tag{20}$$

In the CLPM, the factor CLPM, the RI-CLPM, the STARTS model, and the LCM-SR, accumulating factors are not used, meaning that $\left\{A_{xi} + (t-1)B_{xi}\right\}$ and $\left\{A_{yi} + (t-1)B_{yi}\right\}$ are set to 0. In the ALT model, all parts of the dynamic equations are used. In the LCS model, A_{xi} and A_{xi} are included, while $B_{xi} = 0$ and $B_{yi} = 0$. Moreover, in many applications, the dynamic residuals are also fixed to zero (i.e. $d_{xit} = 0$ and $d_{yit} = 0$) for estimation purposes.

Because all terms in the dynamic equations contribute to the lagged predictors, their influences feed forward through the lagged relations and accumulate at later time points. This implies that the accumulating factors A and B, and the dynamic residuals d have direct and indirect effects on f^*, and thus on the observed scores.

Specification of longitudinal cross-lagged models based on the unified framework

Using the unified framework presented above we can easily see that there are seven components that may or may not be included in the model. These components are:

unique factors ϵ (measurement equations), group means μ, random intercepts I, random slopes S (decomposition equations), accumulating factors A and B, and dynamic residual d (dynamic equations). In Table 2 we provide an overview of the seven models we have discussed above, and indicate which of these seven components they include. The corresponding mathematical expressions of the seven models in terms of the unified framework are presented in Table 3. A conceptual diagram to further clarify the relations between the cross-lagged models is provided in Figure 2. Note that we assumed time-invariant autoregressive and cross-lagged parameters in all the models here, though this assumption can be relaxed.

Table 2 provides a useful summary of the distinctive features of the different types of longitudinal cross-lagged models. In addition, it can be also used to consider nested relation between models: When a model contains all but one of the components another model includes, it forms a special case of the latter. Some of these nested relations are easy to recognize and have already been discussed before, such as the nesting of the CLPM under the factor CLPM and the RI-CLPM, and the nesting of all the three models under the STARTS model.

From the table, we can also see other nested relations. For example, if we extend the ALT model with unique factor and free factor loadings (i.e. $A_{xi} + \alpha_{xt}B_{xi}$ and $A_{yi} + \alpha_{yt}B_{yi}$), this general version of the ALT model is reduced to the TCS model (i.e. the general version of the LCS model) by setting the accumulating factors A to zero. By considering the nested relation between the TCS model and the factor CLPM (i.e. the TCS model is reduced to the factor CLPM by setting the accumulating factors to constant values across individuals; Usami et al., 2015), we can also find the close relation between the ALT model and the CLPM: The ALT model with free factor loadings is reduced to the CLPM by setting accumulating factor A to zero and B to constant values across individuals.

Table 2 also clarifies why two seemingly similar models are *not* nested. For instance, the RI-CLPM and the LCM-SR are not nested, as the RI-CLPM contains the component μ_t, which is not included in the LCM-SR, while the latter contains the component S, which is not included in the RI-CLPM. But when the group means μ_t in the RI-CLPM are constrained to be equal over time (i.e. $\mu_t = \mu$), we can estimate it using the mean of the common factor I, while setting S to zero; hence, in the specific case that the group means do not vary over time, the RI-CLPM is a special case of the LCM-SR (i.e. S=0). Alternatively, when there are no individual differences in the slope of the LCM-SR (i.e. variances of slope factors S are zero), this model actually becomes a special case of the RI-CLPM, as it is based on the assumption that the group means follow a linear trend, while in the RI-CLPM no such constraint is imposed on the group means.

Because all the cross-lagged models we have discussed can be considered as special cases of the unified framework, the reader might assume that we want to advocate this unified model as the most appropriate model to start with. However, this is *not* actually possible since the unified model would be unidentified due to over-parameterization. Instead, our goal for introducing this framework is to provide a general structure that helps to relate the many diverse modeling approaches that are currently in use and that focus on the cross-lagged relations between variables.

Interpretation of the cross-lagged parameters and choosing between the models

Although previous studies proposed different longitudinal models to examine reciprocal effects, the interpretive differences in the cross-lagged parameters (γ) have rarely been discussed explicitly in the literature. The unified framework we present here, clarifies that the interpretations can largely differ depending on what factors are used in specifying a cross-lagged model.

As indicated before, the first category consists of the models that do not explicitly

account for developmental trajectories (the CLPM, the factor CLPM, the RI-CLPM and the STARTS model). As such, if there are individual differences in developmental changes over time, these are not separated from the reciprocal process that is modeled with cross-lagged relations. In addition, the CLPM and the factor CLPM do not account for stable individual differences; hence, their cross-lagged relations are also affected by stable between-person differences. This is not a major problem when one wants to investigate how individual differences in one variable can predict the change in individual differences in another variable, but researchers should refrain from making claims beyond such descriptive statements. The RI-CLPM and the STARTS model account for stable between-person differences through the inclusion of random intercepts I, such that the cross-lagged relations only capture within-person changes after the individuals$'$ expected scores (captured by the sum of the temporal group means μ_t and the individual deviation from these means I) are accounted for. Given that most causal relations addressed in behavioral science research can be argued to operate at the within-person level, applying the RI-CLPM and the STARTS seems to be a reasonable choice if the primary analytic purpose is to infer reciprocal relation between variables. However, just separating the within-person cross-lagged parameters from stable between-person differences does not fully justify a causal interpretation (e.g. Murayama et al., 2017): Whilst this approach controls for omitted time-invariant variables (i.e., trait factor I), it does not control for omitted time-varying variables that are serially correlated. Thus, including some time-varying covariates in the model might help researchers make a stronger arguments about causality when applying the RI-CLPM and the STARTS model.

When considering the second category of models (the LCM-SR, the ALT model, and the LCS model), the defining feature of these models is that they specifically aim to model the group-level developmental trajectory and the individual differences in these trajectories. However, these models achieve this in different ways. The LCM-SR, which

clearly separates the individual developmental trajectories from the reciprocal effects, may at first seem like the ideal solution if the primary analytic purpose is to infer the reciprocal relation or causality between variables. However, there are costs to this approach. Specifically, the LCM-SR decomposes the scores into one part due to individual trends described by growth factors (I and S) and another part that represents the individual′s temporal deviation from this trend. The cross-lagged relations are modeled between the latter residual components, thus, the cross-lagged relations only relate to dynamics $after$ individual developmental changes have already been accounted for by the slope factors S. This may well be a case of throwing the baby out with the bathwater, if the key interest is actually in the reciprocal relations between developmental processes. One could say that the LCM-SR controls for omitted time-varying variables that behave as linear trends and that are the driving force behind the developmental trajectories. However, if the development in one of the observed variables is actually (in part) the result of development in another observed variable, using this LCM-SR is likely to result in biased estimates of reciprocal effects, even though they can provide an inferential basis for within-person processes.

These features of the LCM-SR can be contrasted with the RI-CLPM or the STARTS model: Both RI-CLPM and the STARTS model allow for a group-level trajectory (expressed as μ_t) that can take on any shape, and each individual deviates from this trajectory by the distance of a *constant* (i.e. I). The reciprocal effects are then modeled between the residuals, that is, the deviations from the expected scores based on the group trajectory and the individual constant, without controlling for individual developmental changes by the slope factors S. Hence, in this approach, individual differences in development remain in the reciprocal part of the model. However, if individual differences in development are caused by an omitted time-varying variable, this may lead to biased estimates and erroneous conclusions about the reciprocal relations. In

other words, the choice between the LCM-SR and the RI-CLPM/STARTS model lie in what aspect is considered a critical component of reciprocal effects. If individual differences in the developmental trajectories are considered a critical component of reciprocal effects, the RI-CLPM or the STARTS model might be an appropriate choice. On the other hand, if researchers consider developmental trajectories to be caused by (serially-correlated) omitted time-varying variables (that exhibit linear trends), the (linear) LCM-SR might be a better choice.

Accumulating factors (A and B), which are included in the ALT and the LCS models, do not indicate stable individual differences, because their influences accumulate through the dynamics of the process over time. In these models, the lagged relations are not separated from the accumulating factors, and as a result the accumulating factors have both direct and indirect effects on the observed scores. The critical point is that two primary analytic purposes in applying cross-lagged models (i.e. inferring reciprocal effects between variables, and describing latent trajectories of individuals) are intertwined in these models, such that the lagged relations represent both the short-term dynamics and the long-term developmental trajectories. While this may be a reasonable assumption for some instances, it may be inappropriate for other instances, potentially leading to erroneous conclusions. More specifically, whilst the inclusion of accumulating factors can provide the inferential bases for within-person processes, which would also allow us to control for omitted time-varying variables, there is a risk that accumulating factors wrongly account for individual differences in developmental changes, resulting in biased estimates of reciprocal effects if such individual differences are considered to be constituent components of reciprocal effects.

It should be noted that the lagged parameters from the LCM-SR and the ALT model are identical under certain circumstances (cf. Hamaker, 2005), which suggests that in some scenarios, describing the short-term and long-term fluctuations with the same set

of lagged relations as the ALT model does, is arguably equivalent to separating the developmental trajectory from the dynamic residuals as the LCM-SR does. This is also consistent with results reported by Wang and Maxwell (2015), who focused on the effect of x on y in longitudinal multilevel models in the absence of lagged relations. They found that detrending x, detrending x and y, or including time as a covariate all led to (very) similar results regarding the effect of x on y.

A formal analysis of causal inference in cross-lagged models

To further clarify the causal implications of the cross-lagged models that we have discussed so far, here we provide a formal discussion of causal inference in these cross-lagged models. Specifically, we will place these models within the potential outcome (or counterfactual) approach (Rubin Causal Model; Rubin, 1974), which is currently considered the standard framework for defining causal effects (e.g. Imai, Keele & Tingley, 2010 and Huang & Yuan, 2017 for mediation analysis).

In this framework, the potential outcome for variable Y is denoted as $Y_i(x_i)$ for individual i (Huang & Yuan, 2017; Robins & Hernan, 2009): It is the outcome of Y that we would have observed for this individual if $X_i = x_i$. If X is a dichotomous treatment in cross-sectional and experimental designs, with 0 indicating the control group condition, and 1 indicating the treatment group condition, each individual has two potential outcomes: $Y_i(0)$ and $Y_i(1)$. It is assumed that the potential outcomes of individual i are unaffected by the particular assignment of treatments to the other individuals, and that the potential outcomes of individual i are assumed not to be affected by the assignment of treatment to the same individual i . Then, the individual level effect of the treatment on the outcome is defined as $Y_i(1) - Y_i(0)$, and the average causal effects of the treatment on the outcome is defined as $E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$ over the study population. If there are multiple levels of X (e.g. if X is a continuous variable), the individual level

effects when X increases one unit from the reference value x^r is defined as $Y_i(x_i^r + 1) - Y_i(x_i^r)$, and the average causal effect in this case is defined as $E[Y(x_i^r + 1) - Y(x_i^r)] = E[Y(x_i^r + 1)] - E[Y(x_i^r)]$ over the study population.

In general, these causal effects are not directly estimable from the observed data. Instead, to estimate them, assumptions are needed that link the potential outcomes to the observed values (Huang & Yuan, 2017). Below, we begin with discussing the identifiability assumptions explained by Huang and Yuan (2017) and Robins and Hernan (2009) in the context of longitudinal designs, and we show how these lead to the conclusion that the cross-lagged parameters from the traditional CLPM represents the average causal effects. However, these original identifiability assumptions are rather restrictive and do not allow for unobserved confounding. Therefore, we discuss how these assumptions can be relaxed in order to account for other unobserved confounding due to time-invariant factors, and show how this leads to the conclusion that the cross-lagged parameters from the RI-CLPM and the STARTS model reflect the average causal effects. We also discuss the possibilities of other cross-lagged models to uncover causal effects, and caution that model-specific assumptions, which might be unrealistic in practice, have to be met in order for cross-lagged parameters in each model to be the causal effect.

Identifiability assumptions. Here, we assume that X and Y are variables that are observed longitudinally, with $Y_{it}(x_{i(t-1)})$ being the potential outcome that denotes the outcome of variable Y at time point t that we would have observed for individual i if $X_{i(t-1)} = x_{i(t-1)}$ at the previous time point $(t-1)$. Furthermore, let $\boldsymbol{L_{i(t-1)}} = (Y_{i(t-1)}, \boldsymbol{L_{i(t-1)}^{-\prime}})'$ be a vector of observed variables that includes the time-varying confounder $Y_{i(t-1)}$, which is controlled for via the autoregressive effect, as well as other observed time-varying confounders $\boldsymbol{L_{i(t-1)}^{-}}$ at $t-1$. Additionally, $\boldsymbol{U_{i(t-1)}} = (I_{yi}, \boldsymbol{U_{i(t-1)}^{-\prime}})'$ is a vector of unobserved variables that includes the time-invariant stable trait factor score I_{yi}, which represents a composite of all

time-invariant confounders, and $\boldsymbol{U}^{-}_{\boldsymbol{i(t-1)}}$, which represents a set of time-varying confounders at $t-1$.

As illustrated by Huang and Yuan (2017) and Robins and Hernan (2009) in a longitudinal design, under the following identifiability assumptions we can show that the cross-lagged parameters from the CLPM (i.e. γ_{yt}) can have a causal interpretation:

Assumption 1 (consistency): If $X_{i(t-1)} = x_{i(t-1)}$ for individual i, then $Y_{it}(x_{i(t-1)}) = Y_{it}$ for that individual.

Assumption 2 (no unobserved confounders): $Y_{it}(x_{i(t-1)}) \perp X_{i(t-1)} | \boldsymbol{L_{i(t-1)}} = \boldsymbol{l_{i(t-1)}}$ for each possible value x of X and l of L.

Assumption 3 (positivity): If $Pr(\boldsymbol{L_{it}} = \boldsymbol{l_{it}}) \neq 0$, $Pr(X_{it} = x_{it} | \boldsymbol{L_{it}} = \boldsymbol{l_{it}}) > 0$ for all x.

These first two assumptions are standard in causal inference (e.g. Huang & Yuan, 2017), whereas the third one is required when evaluating the weights posited to adjust observed confounders (e.g. Robins & Hernan, 2009), which we will briefly discuss at the end of this section.

Assumption 1 is known as consistency, and indicates that the observed outcome Y_{it} is equal to the potential outcome $Y_{it}(x_{i(t-1)})$ if observed level of X at the previous time point is $X_{i(t-1)} = x_{i(t-1)}$. An implicit assumption here is that measurement error is not present. Another implicit assumption is the Markov property to achieve consistency for the potential outcome: The current potential outcome of Y only depends on the previous state of X. In other words, there is no "carryover" effect (Imai & Kim, under review) of past values prior to $t-1$. Assumption 2 is known as no unobserved confounders, and indicates that there is no unobserved confounding between the current potential outcome of Y and the previous state of X after accounting for (the previous state of) the observed

confounders $\boldsymbol{L}$.

Together, these first two assumptions ensure that $E_i(Y_{it}|X_{i(t-1)} = x_{i(t-1)}, \boldsymbol{L_{i(t-1)}} = \boldsymbol{l_{i(t-1)}})$ is identifiable from the observed data and model parameters. Suppose that X and $\boldsymbol{L}$ are all linearly related to Y (as is the case for the relation between X and Y in the cross-lagged models we have considered); then, the average potential outcome can be expressed as

$$\begin{aligned} E_i(Y_{it}(x_{i(t-1)})) &= E_i(Y_{it}|X_{i(t-1)} = x_{i(t-1)}, \boldsymbol{L_{i(t-1)}} = \boldsymbol{l_{i(t-1)}}) \\ &= c_{yt} + \beta_{yt} y_{i(t-1)} + \gamma_{yt} x_{i(t-1)} + \boldsymbol{l'^{-}_{i(t-1)}} \boldsymbol{\delta_{yt}}, \ \ \forall t (= 2, 3, \ldots, T) \end{aligned} \tag{21}$$

Where c_{yt} is an intercept and β_{yt}, γ_{yt} and $\boldsymbol{\delta_{yt}}$ are regression coefficients. If we use time centered valuables such as $y^*_{i(t-1)} = y_{i(t-1)} - \mu_{y(t-1)}$, $x^*_{i(t-1)} = x_{i(t-1)} - \mu_{x(t-1)}$ and $\boldsymbol{l^{-*}_{i(t-1)}} = \boldsymbol{l^{-}_{i(t-1)}} - \boldsymbol{\mu_{l(t-1)}}$ at each time point, alternative formulation can be obtained as

$$E_i(Y_{it}(x_{i(t-1)})) = \mu_{yt} + \beta_{yt} y^*_{i(t-1)} + \gamma_{yt} x^*_{i(t-1)} + \boldsymbol{l'^{-*}_{i(t-1)}} \boldsymbol{\delta_{yt}}, \tag{22}$$

where $\mu_{yt} = c_{yt} + \beta_{yt}\mu_{y(t-1)} + \gamma_{yt}\mu_{x(t-1)} + \boldsymbol{\mu'_{l(t-1)}}\boldsymbol{\delta_{yt}}$.

Obviously, this expression is mathematically equivalent to the (conditional) expected outcome of Y_{it} from the CLPM that assumes observed confounders. Using this expression, the average causal effect when X increases one unit from the reference value x^r at time $t-1$ is given by

$$\begin{aligned} E_i(Y_{it}(x^r_{i(t-1)} + 1)) - E_i(Y_{it}(x^r_{i(t-1)})) &= [\mu_{yt} + \beta_{yt} y^*_{i(t-1)} + \gamma_{yt}(x^{*r}_{i(t-1)} + 1) + \boldsymbol{l'^{-*}_{i(t-1)}} \boldsymbol{\delta_{yt}}] \\ &\quad - [\mu_{yt} + \beta_{yt} y^*_{i(t-1)} + \gamma_{yt} x^{*r}_{i(t-1)} + \boldsymbol{l'^{-*}_{i(t-1)}} \boldsymbol{\delta_{yt}}] \\ &= \gamma_{yt}, \end{aligned} \tag{23}$$

where $x^{*r}_{i(t-1)} = x^r_{i(t-1)} - \mu_{x(t-1)}$. Thus, if the true data generating model satisfies the identifiability assumptions (i.e. consistency, no unobserved confounders) and two additional ones (i.e. X and $\boldsymbol{L}$ are linearly related to potential outcomes, normal

residuals), γ_{yt} in the CLPM (with observed confounders $\boldsymbol{L}^{-}$) represents the average causal effect across individuals from x at time $t-1$ to y at time t. Likewise, it can be shown that γ_{xt} also represents the average causal effect from y at time $t-1$ to x at time t.

However, if unobserved confounders $\boldsymbol{U_{it}}$ are present as the cause of X_{it} as well as Y_{it} in the true data generating model, this implies that $E_i(Y_{it}(x_{i(t-1)})) \neq E_i(Y_{it}|X_{i(t-1)} = x_{i(t-1)}, \boldsymbol{L_{i(t-1)} = l_{i(t-1)}})$, and thus γ_{yt} in the CLPM does not represent the (average) causal effect any more, because this model fails to account for $\boldsymbol{U_{it}}$ (i.e. unobserved confounding). Unfortunately, unobserved confounding is very likely in observation studies. The major weakness of observation studies is that, unlike in randomized experiments with full compliance, the above identifiability conditions are not guaranteed by design (Robins & Hernan, 2009). An additional problem is that the presence of unobserved confounding can not be empirically tested.

Extending identifiability assumptions to account for time-invariant unobserved confounding. To overcome the problem of unobserved confounding, accounting for a (time-invariant) stable trait factor I, as the RI-CLPM and the STARTS model do, can mitigate the potential risk. Specifically, if we relax the assumptions of the second and third identifiability assumptions by allowing the presence of stable trait factor I into the true data generating model as

Assumption 2': (no unobserved confounders after controlling for a stable trait factor) $Y_{it}(x_{i(t-1)}) \perp X_{i(t-1)}|\boldsymbol{L_{i(t-1)} = l_{i(t-1)}}, I_{yi}$

Assumption 3': (positivity with controlling for a stable trait factor) If $Pr(\boldsymbol{L_{it} = l_{it}}|I_{yi}) \neq 0$, $Pr(X_{it} = x_{it}|\boldsymbol{L_{it} = l_{it}}, I_{yi}) > 0$ for all x.

then, $E_i(Y_{it}(x_{i(t-1)})) = E_i(Y_{it}|X_{i(t-1)} = x_{i(t-1)}, \boldsymbol{L_{i(t-1)} = l_{i(t-1)}}, I_{yi})$. Suppose that X, $\boldsymbol{L}$

and I are all linearly related to Y, the average of potential outcomes can be expressed as

$$E_i(Y_{it}(x_{i(t-1)})) = c_{yt} + I_{yi} + \beta_{yt} y_{i(t-1)} + \gamma_{yt} x_{i(t-1)} + \boldsymbol{l}'^{-}_{\boldsymbol{i(t-1)}} \boldsymbol{\delta_{yt}} \qquad \forall t \quad (= 2, 3, \ldots, T). \tag{24}$$

We can rewrite this equation as

$$\begin{aligned} E_i(Y_{it}(x_{i(t-1)})) &= [c_{yt} + I_{yi} + \beta_{yt}(I_{yi} + \mu_{y(t-1)}) + \gamma_{yt}(I_{xi} + \mu_{x(t-1)}) + (\boldsymbol{I_{li}} + \boldsymbol{\mu_{l(t-1)}})' \boldsymbol{\delta_{yt}}] \\ &\quad + \beta_{yt} y^*_{i(t-1)} + \gamma_{yt} x^*_{(t-1)} + \boldsymbol{l}'^{*-}_{\boldsymbol{i(t-1)}} \boldsymbol{\delta_{yt}} \\ &= [\mu_{yt} + I^*_{yi}] + \beta_{yt} y^*_{i(t-1)} + \gamma_{yt} x^*_{(t-1)} + \boldsymbol{l}'^{*-}_{\boldsymbol{i(t-1)}} \boldsymbol{\delta_{yt}}, \end{aligned} \tag{25}$$

where $I^*_{yi} = I_{yi} + \beta_{yt} I_{yi} + \gamma_{yt} I_{xi} + \boldsymbol{I'_{li}} \boldsymbol{\delta_{yt}}$, $\mu_{yt} = c_{yt} + \beta_{yt} \mu_{y(t-1)} + \gamma_{yt} \mu_{x(t-1)} + \boldsymbol{\mu'_{l(t-1)}} \boldsymbol{\delta_{yt}}$, $y^*_{i(t-1)} = y_{i(t-1)} - (I_{yi} + \mu_{y(t-1)})$, $x^*_{i(t-1)} = x_{i(t-1)} - (I_{xi} + \mu_{x(t-1)})$ and $\boldsymbol{l^*_{i(t-1)}} = \boldsymbol{l_{i(t-1)}} - (\boldsymbol{I_{li}} + \boldsymbol{\mu_{l(t-1)}})$. $\boldsymbol{I_{li}}$ is a vector of stable component of observed confounders for an individual i. If stationarity of β_{yt}, γ_{yt} and δ_{yt} is satisfied, $I^*_{yi} = I_{yi} + \beta_y I_{yi} + \gamma_y I_{xi} + \boldsymbol{I'_{li}} \boldsymbol{\delta_y}$ has a time-invariant feature, indicating that in this case Equation 25 is mathematically equivalent to the (conditional) expected outcome of Y_{it} from the RI-CLPM that assumes observed confounders $\boldsymbol{L}^-$.

From Equation 25, under a weaker identifiability assumption (i.e. no unobserved confounders after controlling for a stable trait factor) and other assumptions (i.e. X, $\boldsymbol{L}$ and I are linearly related to potential outcomes, normal residuals), the average causal effect can be expressed as

$$E_i(Y_{it}(x^r_{i(t-1)} + 1)) - E_i(Y_{it}(x^r_{i(t-1)})) = \gamma_{yt}. \tag{26}$$

This result indicates that under the weaker identifiability assumption γ_{yt} in the RI-CLPM (with stationarity assumption and observed confounders) represents the (average) causal effect, while γ_{yt} in the CLPM does not because this model fails to account for time-invariant unobserved confounding variables (i.e. stable trait factor I). In other words, the RI-CLPM (with stationarity assumption and observed confounders) can provide a

stronger basis to infer causal effects by allowing for the presence of time-invariant unobserved confounding variables in estimating γ. Because the stable trait factor is likely to exist in many situations in psychology (i.e. most variables in psychology have individual differences in mean scores over time), the RI-CLPM can be considered as a reasonable approach that can infer (average) causal effects more safely, compared to the CLPM.

In the same way, it can be shown that the STARTS model (with stationarity assumption and observed confounders), which accounts for the presence of unique factors (i.e. measurement errors as well), can provide a stronger basis than the RI-CLPM to infer causal effects by loosening assumption 1.

Time-varying unobserved confounding and model selection. One may consider that other longitudinal models such as the LCM-SR, the ALT model and the LCS model, provide even stronger inferential basis for causal effects than the RI-CLPM and the STARTS model, because these models implicitly allow for other types of unobserved confounding variables. As we have discussed, the ALT and the LCS models control for time-varying unobserved confounding variables (accumulating factors A and B) that have both direct and indirect effects on the observed scores through lagged relations. Also, the LCM-SR controls for time-varying unobserved confounding variables that behave as linear trends (S). In particular, in the ALT and the LCS models indirect effects are expressed by (the linear sum of) complex functions of autoregressive parameters and cross-lagged parameters which are accumulated over time (see Equation 17 for occasion $t = 3$ in the LCS model). If unobserved confounders behave in such a way in true data generating model, cross-lagged parameters in these cross-lagged models represent causal effects, whereas those in the CLPM and the RI-CLPM do not.

More specifically, from the similar procedures, it can be shown that the ALT and the LCS models (with observed confounders) can provide a stronger basis than the RI-CLPM and the STARTS model to infer causal effects by loosening the assumption 2'

(i.e. no unobserved confounders after controlling for a stable trait factor), whereas the LCM-SR demands a stationarity assumption as well as observed confounders for γ to represent causal effects. However, when influences of time-varying unobserved variables do not change linearly over time (i.e. the LCM-SR) or in accumulative ways (i.e. the LCS and the ALT models), these models cannot provide accurate estimates of causal effects.[4]

Hence, we can conclude that no single cross-lagged model is ever superior to the others for estimating causal effects, and all models (except for the CLPM and the factor CLPM) can speak to causality if they model unobserved confounders correctly when they are present. In other words, the crucial point for choosing the cross-lagged model is whether the model provides adequate control of time-varying and time-invariant confounders. To judge this, substantive knowledge about potential confounders and their characteristics (e.g. are they time-invariant or time-varying? do their means change linearly or non-linearly?) that researchers are expecting is important. In addition, fit indices should be a useful tool to assess the potential misspecification.

In conclusion, if researchers are interested in causal inference, model selection should be made based on the balance between substantial knowledge about potential confounders and their characteristics, and the bottom up assessment of potential misspecification using model fit indices. If researchers do not have good knowledge about confounders, controlling for at least a stable trait factor might be a safe choice as long as the fit is good enough.

Marginal structural modeling. Regardless of the choice of the cross-lagged models, in order for γ to be the causal effect, model-specific assumptions (e.g., distributional assumptions in unique factors, no time-varying unobserved confounders, structure of change in time-varying unobserved confounder, and linear relation between (bivariate) outcomes) have to be met but these might be unrealistic in practice. However, recent statistical developments have made it possible to relax some of these assumptions.

Specifically, marginal structural models (MSMs; Robins, 1999; Robins, Hernan, & Brumback, 2000) do not demand researchers to model the particular relation between a (potential) outcome and (observed) confounding variables, whereas a linear relation was implicitly assumed in the previous discussion (i.e. Equations 22 and 25) . Since its introduction by Robins (1999), MSMs have rapidly gained popularity among applied researchers in biomedical and other fields (e.g. VanderWeele, Hawkley, Thisted, & Cacioppo, 2011 in psychology) as a tool for causal inference from longitudinal data in observational studies (Imai & Ratkovic, 2015). If the three identifiability conditions hold, in MSMs, instead of including observed confounding variables to explain (potential) outcomes as in Equation 22, a propensity score for each individual with $X_{i(t-1)} = x_{i(t-1)}$ is first estimated from the observed confounding variables using some sort of function (e.g. logistic regression) at each time point. Then, the influence of these confounding variables on outcomes Y_{it} is adjusted and parameters are estimated using inverse probability weights (IPW), which are typically based on the product of a propensity score estimated separately at each time point.

However, it should be noted that in MSMs the outcome at a particular measurement occasion (typically the last) to be their focus; hence, it is not really about obtaining the reciprocal dynamics between two variables over time, as cross-lagged models we have discussed. Therefore, MSMs are applicable only when researchers aim to investigate causal effects through unidirectional relations of variables (See VanderWeele et al., 2011 for example). Another critical issue is that there are some computational challenges to precisely estimate the IPW (e.g. the IPW is highly sensitive to misspecification of the function for calculating propensity scores even when the number of time points is moderate; Imai & Ratkovic, 2015). There are several recent developments in this field to overcome this problem (e.g. Bang & Robins, 2005 for doubly robust estimators; Imai & Ratkovic, 2015 for covariate balancing propensity score (CBPS)

methodology). The possible dependence between confounders and measurements might also be a challenging issue. For example, if the time-varying confounders are affected by the time-varying exposure (and the confounders, in turn, affect the exposure) controlling for these confounders in longitudinal models could become difficult.

An Example Using Empirical Data

Here, we present an example using empirical data to show how parameter estimates for the cross-lagged effects are different depending on the choice of the model (see also Usami et al., 2015; Bainter & Howard, 2016). This example investigates the reciprocal relation between adolescents′ perceived exposure to smoking in movies (X) and their smoking intensity (Y), using data from the Minnesota Adolescent Community Cohort (MACC) Study 2000-2013. The MACC Study is a prospective cohort study designed to expand the understanding of the transitional process from non-smoking to smoking during adolescence and to examine the effect of state- and local-level tobacco prevention and control programmes on youth in Minnesota (Choi, Forster, Erickson, Lazovich, & Southwell, 2012). For illustrative purposes we used a sample of 4,671 adolescents who were surveyed from ages 15 to 20 every six months in most of the years of the survey. When participants joined two surveys in a year, only the response data from the first survey were used to construct the dataset of $T = 6$, which is common practice when analysing longitudinal data. More detailed information about the study design and the population of MACC is available in Choi et al (2012) and on the website of the Inter-university Consortium for Political and Social Research (https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36282) [5].

Perceived exposure to smoking in movies and smoking intensity were assessed during each round of data collection. Participants were asked to report how often they saw actors and actresses smoking when they watched movies, with four response options:

most of the time (4), some of the time (3), hardly ever (2) and never (1). This measure assessed participants′ exposure to smoking in movies and their recognition of these images, similar to the construct of receptivity to tobacco marketing (Choi et al., 2012). The data also included an index of smoking intensity, which represents how frequently the participants smoked (see Choi et al., 2012, for details). There were two sources of missing data: missing by design and attrition. We used all available data in estimating parameters of cross-lagged models (we removed data from one participant who was missing for all of the variables in all time points). All analyses were conducted using the lavaan package (Rosseel, 2012) in R (R Core Team, 2016) using the full information maximum likelihood estimation method. The lavaan source code is available in the Online Supplemental Materials.

Table 4 shows the parameter estimates and fit indices from various models we discussed in the paper. Although we assumed stationarity (e.g. equality of autoregressive effects, cross-lagged effects, unique variances and covariances over time), we can actually loosen these constraints if necessary. The assumption of stationarity is typically made in the application of the LCS model, but for other models this assumption is less common. The STARTS model and the ALT model both resulted in an improper solution; as such Table 4 only provides the estimates for the CLPM, the factor CLPM, the RI-CLPM, the LCM-SR and the LCS model [6]. The CLPM, which is most widely used in applied research, showed statistically significant and positive cross-lagged estimates (perceived exposure to smoking in movies predicted the later smoking intensity, and vice versa), but also had the worst fit among the models. The factor CLPM and the RI-CLPM fitted better, and resulted in non-significant estimates for both cross-lagged parameters in the RI-CLPM, while the factor CLPM showed statistically significant cross-lagged estimates for γ_y (meaning that more perceived exposure to smoking in movies predicted stronger smoking intensity).

The LCS model fitted better than these models, and resulted in statistically significant and *negative* cross-lagged estimates for γ_x (meaning that stronger smoking intensity predicted less perceived exposure to smoking in movies). It should be noted that the LCS model showed relatively larger standard errors in comparison to other models. This may suggest instability of parameter estimates due to possible multicollinearity caused by large correlations between accumulating factors (A) and latent true scores at the previous time point (Usami et al., 2015). Finally, the LCM-SR model had the best model fit, and resulted in non-significant estimates for both cross-lagged parameters, similar to the RI-CLPM.

This example clearly demonstrates the risk of drawing different conclusions based on the cross-lagged parameters from different models: We obtained a significant negative cross-lagged parameter (the LCS model), significant positive cross-lagged parameters (the CLPM and the factor CLPM), and non-significant cross-lagged parameters (the RI-CLPM and the LCM-SR). By focusing on model fit from the view of model misspecification the LCM-SR would be the most appropriate model for these data. Moreover, since the model fit of the LCM-SR is so much better than that of the RI-CLPM (e.g. the difference in AIC is more than 200, and this is also true for the difference in BIC, which could be considered as overwhelming evidence for the LCM-SR), it would seem safe to conclude that there are omitted time-varying variables that affect both observed variables. If researchers are interested in causal inferences and they can reasonably expect unobserved time-varying confounders that change linearly (as an S factor), choosing the estimation results from the LCM-SR seems to be reasonable. In this case, it can be concluded that there is little evidence for causal effects.

The RI-CLPM could be an option if one aims to infer reciprocal predictive relations between variables without being certain that there are unobserved time-varying confounders that change linearly, and individual differences in the developmental changes

of smoking intensity and perceived exposure to smoking in movies are considered to be critical components of reciprocal effects. In this case, it would be reasonable to conclude that there are no significant reciprocal effects between these variables. However, we need to consider that their estimates of reciprocal effects might be biased due to the (serially-correlated) omitted time-varying variables.

One important note is that such analytic purposes are inherently inseparable from another major analytic purpose of using a cross-lagged model: estimating developmental trajectories. Even if one is not interested in developmental trajectories, the presence of certain types of unobserved confounders sometimes demands the correct specification of developmental trajectories to accurately estimate causal effects. Conversely, if one is interested in developmental trajectories but individual differences in the developmental changes are considered to be critical components of causal effects rather than (unobserved) confounders, overly controlled common factors in the model are likely to cause biased estimates for causal effects.

General Discussion

Inferring reciprocal effects between variables is a central aim of behavioral research. To address reciprocal effects, a variety of longitudinal models that include cross-lagged relations have been proposed. However, as these models were proposed in diverse contexts with different backgrounds, the similarities and differences between these longitudinal cross-lagged models have been unclear, making it difficult for researchers to select a model that suits the goal of their research. To facilitate the understanding of the strengths and weakness of these models and to promote informed model selection, we have proposed a unified framework that helps to clarify the conceptual and statistical differences between various longitudinal cross-lagged models (i.e. the CLPM, the factor CLPM, the RI-CLPM, the STARTS model, the LCM-SR, the ALT model, and the LCS model). The

unified framework effectively summarizes the relation between these models, mainly based on whether and how unique factors, different types of common factors, and developmental trajectories are specified.

We suggest that, when investigating cross-lagged relations, researchers should select a model based on the analytic purpose (i.e. investigating reciprocal relations or causality) and substantial knowledge (i.e. what types of confounders researchers are expecting) as well as the result of bottom up assessment of potential misspecification using model fit indices. This combination of theory and empirically driven approaches would help researchers to make the most informed decision when selecting models for analysis.

Currently, it is common that researchers use the CLPM as a default choice without considering any other alternative models to test reciprocal effects, even though these alternatives may provide better insight into their data. For example, it has been shown that reliance on the CLPM alone can potentially lead to erroneous conclusions regarding the presence, predominance, and sign of reciprocal effects (Hamaker et al., 2015) .

We have argued that model comparison is an important step toward better understanding of the reciprocal relations (or causality) between variables, but we are also aware of some practical issues which may be worth considering. First, different models require different numbers of time points for model identification, as we summarized in Table 1. It shows that the CLPM is the only model that is identified when $T = 2$ or more, and the required number of time points increases as the model becomes more complex. The ALT model requires the largest number of time points ($T = 4$ or more when assumption of stationarity of parameters is made, and $T = 5$ or more when this assumption is relaxed) mainly because it has more than one accumulating factor (A and B). These observations indicate that, although model comparison should be encouraged, it is often the case that the data researchers acquire do not allow for such model comparisons due to the limited number of time points. In fact, Hamaker et al. (2015)

found that 45% of the studies that used the CLPM had only two waves of data. Hence, prior to data collection, researchers should consider carefully how many time points they should record depending on which cross-lagged model they will use for analysis.

Second, our empirical example highlighted the problem of improper solutions when comparing different types of models. Although not well documented in the literature (although see, Cole, 2005), we have frequently heard from applied researchers that they encounter improper solutions when they applied these different types of longitudinal models (other than the CLPM). An inherent problem is that improper solutions do not always mean that the model is misspecified. A correctly specified model could produce improper solutions due to sampling error, especially when the model is complex, which is the case for the longitudinal models we reviewed. When the model contains unique factors (e.g. the STARTS model or the LCS models), it may be helpful to use multiple indicators (rather than a single indicator). Cole et al. (2005) found that the TSO model, which assumes multiple indicators in the STARTS model (but no errors are assumed in common factors for indicators), encountered fewer improper solutions. This indicates that the inclusion of multiple indicators may help to reduce the risk of improper solutions. In addition, a recent study (L*ü*dtke, Robitzsch & Wagner, 2017) showed that Bayesian estimation can effectively address problems of estimation including improper solutions in the STARTS model by specifying prior distributions that are weakly informative. Further research is needed to develop more strategies to address the issue of improper solutions.

It is also important to note that, although some models are well-suited to examine reciprocal effects (e.g. the RI-CLPM), we still need to consider potential unobserved time-varying confounders to make an accurate causal inference, as these models do not control for them. Such time-varying (and time-invariant) confounders can be easily implemented in the existing longitudinal models. Indeed, Bollen and Brand (2010) proposed a general panel model with lagged effects to model univariate longitudinal data

using both time-varying and time-invariant covariates. In their model, they include a latent time-invariant variable which is the same as an accumulating factor A in the unified framework proposed here.

One limitation of the current manuscript is that we did not discuss models that incorporate the potential random effects other than random intercepts. For example, in comparison to the CLPM, the RI-CLPM incorporates stable trait factors with fixed factor loadings (= 1). These stable trait factors are analogous to random effects in temporal group means. Similarly, the LCS model can be seen as an extension of the CLPM which includes random intercepts in cross-lagged regressions (Usami et al., 2015). These observations indicate that stable trait factors or accumulating factors can serve as random effects of intercepts in cross-lagged models. However, these models do not consider random effects (i.e. individual differences between participants) in other parameters: autoregressive coefficients, cross-lagged (coupling) coefficients, and unique and residual variances. This is a hidden, but strong assumption in the existing longitudinal cross-lagged models, as it is plausible that autoregressive and reciprocal effects are different between individuals. There have been recent developments to combine between- and within-person information in a multilevel framework: For example, Bringmann et al., (2013) applied the network approach based on the multilevel vector autoregression (VAR) model, allowing for individual differences in the lagged parameters (see also Schuurman, Ferrer, de Boer-Sonnenschein, and Hamaker, 2016). These multilevel autoregressive models are used to analyze intensive longitudinal data consisting of 20 or more repeated measurements per person. While allowing for individual differences in the lagged parameters would be of interest in many applications, it is not clear whether this would be feasible with the relatively small number of repeated measures in traditional longitudinal research.

In recent decades, the literature on longitudinal data design and analysis has been rapidly growing. For example, with intensive longitudinal data, Bringmann, Hamaker,

Vigo, Aubert, & Borsboom, (2017) proposed the semi-parametric TV-AR model to capture nonstationary processes that may arise in such data for univariate cases. There is also a growing body of literature on continuous time modeling using SEM (e.g. Deboeck & Preacher, 2016; Voelkle, Oud, Davidov, & Schmidt, 2012), which can effectively address unequally spaced observations and the lag-problem (i.e. the fact that the strength of a lagged relation is influenced by the interval between the observations; see Gollob and Reichardt, 1987). Furthermore, there has been major development in recent classification techniques for longitudinal data such as SEM Trees (Brandmaier, Oertzen, McArdle, & Lindenberger, 2013; Usami, Hayes, & McArdle, 2017) and SEM forests (Brandmaier, Prindle, McArdle & Lindenberger, 2016). These new methodologies expand the scope of the current longitudinal data analysis considerably, but there is still a lot of room for innovation and improvement in how various cross-lagged models can be effectively integrated into these approaches. We hope that our unified framework contributes to a better understanding of the diverse existing longitudinal models and their cross-lagged relations, and helps the integration of the existing cross-lagged models with new approaches.

Allison, P.D. (2005). *Fixed effects regression methods for longitudinal data using SAS.* SAS Institute, Inc, Cary

Bainter, S.A., & Howard, A. L. (2016). Comparing within-person effects from multivariate longitudinal models. *Developmental Psychology, 52*, 1955-1968.

Bang, H., Robins, J.M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics, 61*, 962-973.

Bollen, K.A., & Curran, P.J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research, 32*, 336-383. doi: 10.1177/0049124103260222

Bollen, K.A., & Curran, P.J. (2006). *Latent curve models: A structural equation approach.* Hoboken, NJ: Wiley.

Bollen, K.A., & Brand J.E. (2010). A general panel model with random and fixed effects: a structural equations approach. *Social Forces, 89*, 1-34.

Brandmaier, A.M., Oertzen, T.v., McArdle, J.J., & Lindenberger, U. (2013). Structural equation model trees. *Psychological Methods, 18*, 71-86.

Brandmaier, A.M., Prindle, J,J., McArdle, J.J., & Lindenberger, U. (2016). Theory-guided exploration with structural equation model forests. *Psychological Methods, 21*, 566-582.

Bringmann, L.F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*, e60188, 1-13. doi:10.1371/journal.pone.0060188

Bringmann, L.F., Hamaker, E.L., Vigo, D.E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods, 22,* 409-425.

Chi, E.M., & Reinsel, G.C. (1989). Models for longitudinal data with random effects and AR (1) errors. *Journal of the American Statistical Association, 84*, 452-459.

Cole, D.A., Martin, N.C., & Steiger, J.H. (2005). Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model. *Psychological Methods, 10*, 3-20.

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

Curran, P.J., Howard, A.L., Bainter, S.A., Lane, S.T., & McGinley, J.S. (2013). The separation of between-person and within-person components of individual change: over time A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology, 82*, 879-894.

Deboeck, P.R., & Preacher, K.J. (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling, 23,* 61-75.

Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*, 424-438.

Hamagami, F., & McArdle, J.J. (2001). Advanced studies of individual differences: Linear dynamic models for longitudinal data analysis. In G.A. Marcoulides & R.E. Schumacker (Eds.), New developments and techniques in structural equation modeling. (pp. 203-246). Mahwah, NJ: Erlbaum Associations.

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

Hsiao, C. (2014). *Analysis of Panel Data (3rd ed.)*. London: Cambridge University Press.

Huang, J. & Yuan, Y. (2017). Bayesian dynamic mediation analysis. *Psychological Methods, 22*, 667-686.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods, 15,* 309-334.

Imai, K. and Ratkovic, M. (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association, 110*, 1013-1023.

Imai, K. and Kim, I.S. (under review). When should we use linear fixed effects regression models for causal inference with longitudinal data?

Jongerling, J., & Hamaker, E.L. (2011). On the trajectories of the predetermined ALT Model: What are we really modeling? *Structural Equation Modeling, 18*, 370-382. http://dx.doi.org/10.1037/a0038889

Jöreskog, K.G. (1970). A general method for analysis of covariance structures. *Biometrika, 57,* 239-251. doi: 10.1093/biomet/57.2.239

Jöreskog, K.G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In R.C. Atkinson, D.H. Krantz, R.D. Luce, & P. Suppas (Eds.), *Contemporary developments in mathematical psychology* (pp. 1-56). San Francisco: Freeman.

Jöreskog, K.G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge, MA: Abt Books.

Kenny, D.A., & Zautra, A. (1995). The trait-state-error model for multiwave data. *Journal of Consulting and Clinical Psychology, 63*, 52-59.

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

Lüdtke, O., Robitzsch, A. & Wagner, J. (2017). More stable estimation of the STARTS model: A Bayesian approach using Markov Chain Monte Carlo techniques. *Psychological Methods.* http://dx.doi.org/10.1037/met0000155

Luhmann, M., Schimmack, U., & Eid, M. (2011). Stability and variability in the relationship between subjective well-being and income. *Journal of Research in Personality, 45*, 186-197.

McArdle, J.J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60,* 577-605. doi:10.1146/annurev.psych.60.110707.163612.

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

McArdle, J.J., & Nesselroade, J.R. (2014). *Longitudinal data analysis using structural equation models.* American Psychological Association.

Meredith, W., & Tisak, J. (1984). On "Tuckerizing" curves. Presented at the annual meeting of the psychometric society, Santa Barbara, CA.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107-122. doi:10.1007/BF02294746

Murayama, K., Goetz, T., Malmberg, L-E., Pekrun, R., Tanaka, A., & Martin, A.J. (2017). Within-person analysis in educational psychology: Importance and illustrations. In D.W. Putwain, & K. Smart (Eds.), *British Journal of Educational Psychology Monograph Series II: Psychological Aspects of Education — Current Trends: The role of competence beliefs in teaching and learning.* Oxford: Wiley.

R Core Team (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robins, J. (1999). Marginal structural models versus structural nested models as tools for causal inference. In Halloran, M.E., Berry, D. (eds). *Statistical models in epidemiology: the environment and clinical trials*, New York, NY: Springer, pp.95-134.

Robins, J.M., Hernán, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology, 11*, 550-560.

Robins, J.M. & Hernán, M.A. (2009). Estimation of the causal effect of time-varying exposures. In G. Fitzmaurice, M. Davidian, G. Verbeke & G. Molenberghs (Eds.), *Longitudinal data analysis*. CRC Press.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*, 1-36. URL http://www.jstatsoft.org/v48/i02/

Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688-701.

Schmitt, M.J., & Steyer, R. (1993). The latent state-trait model (not only) for social desirability. *Personality and Individual Differences, 14*, 519-529.

Schuurman, N.K., Houtveen, J.H., & Hamaker, E.L. (2015). Incorporating measurement error in n=1 psychological autoregressive modeling. *Frontiers in Psychology, 6.* doi: 10.3389/fpsyg.2015.01038

Schuurman, N.K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E.L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model.

Psychological Methods, 21, 206-221.

Steyer, R., Schwenkmezger, P., & Auer, A. (1990). The emotional and cognitive components of trait anxiety: A latent state-trait model. *Personality and Individual Differences, 11*, 125-134.

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

Usami, S., Hayes, T., & McArdle, J.J. (2016). Inferring longitudinal relationships between variables: model selection between the latent change score and autoregressive cross-lagged factor models. *Structural Equation Modeling, 23*, 331-342.

Usami, S., Hayes, T., & McArdle, J.J. (2017). Fitting structural equation model trees and latent growth curve mixture models in longitudinal designs: The influence of model misspecification in estimating the number of classes. *Structural Equation Modeling, 24,* 585-598.

VanderWeele, T.J., Hawkley, L.C., Thisted, R.A. & Cacioppo, J.T. (2011). A marginal structural model for loneliness: implications for intervention trials and clinical practice. *Journal of Consulting and Clinical Psychology, 79*, 225-235.

Voelkle, M.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*, 176-192.

Wang, L.P., & Maxwell, S.E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*, 63-83.

Footnotes

[1]The CLPM is also referred to as the autoregressive cross-lagged model, the cross-lagged path model, and the cross-lagged regression model.

[2]Because cross-lagged parameters as well as autoregressive parameters in the RI-CLPM are individual-invariant, *within person* relations captured by this model can be essentially considered as across persons and not within individual. In the general discussion we discuss the extension of models on this point.

[3]For example, Cole et al. (2005) proposed an alternative model called trait-state-occasion (TSO) model. TSO model is a multiple-indicator version of STARTS but no errors are assumed in common factors for indicators, with additional stationarity assumption on autoregressive and cross-lagged parameters (e.g. $\beta_{xt} = \beta_x$). Also, Luhmann, Schimmack, and Eid (2011) tested a modified (bivariate) TSO model, in which factor loadings α_{xt} from stable trait factor scores freely estimated (e.g. use $\alpha_{xt}I_{xi}$ instead of I_{xi}) as well as second-order time-variant autoregressive parameters are added. Although these models are slightly different from each other, all models share the same feature that stable trait factors account for the stable individual differences, allowing us to make an inference about within-person process for their cross-lagged parameters

[4]In the ALT model, there are "predetermined" and "constrained" approaches according to the way of constraining parameters, as we have discussed. Because this difference is irrelevant to the basic assumption of the ALT model (i.e. assuming two accumulating factors), this point is not an important in evaluating the potential of the ALT model to assess causal effects.

[5]Choi et al (2012) investigated the reciprocal relation between adolescents′ perceived exposure to smoking in movies and smoking status (SS). SS is a composite measure and is different from smoking intensity that we used in this example. In addition, Choi et al (2012) focused on the sample of adolescents from ages 12 to 18, and applied only the

CLPM and included second-order autoregressive parameters, while adjusting for gender, race/ethnicity, parent education, level of urbanisation and age cohort. Due to these procedural differences and illustrative purpose of the current example we do not compare analysis results from Choi et al (2012) here.

[6] Many empirical studies have reported the improper solutions in applying the STARTS model, and trait factor variances were negative in the current analysis. Conceptually, one potential reason is that unstable estimates of some parameters (particularly autoregressive parameters) caused by some aspects of the research design (e.g. small sample size) can unduly inflate the variances of the deviation terms (x^*_{it} and y^*_{it}), increasing the risk of obtaining negative estimates of trait factor variances (as well as other variances parameters). In the ALT model, negative variance of intercept factor and singular Hessian matrix were produced. One possible scenario was that true data generating model did not include unobserved time-varying confounders whose behaviors can be expressed by (the linear sum of) the two common factors whose influences accumulate at the later time points, resulting in producing a singular Hessian matrix.

Table 1. The number of parameters and required number of time points for model identification in cross-lagged models.

Model assumption	The number of time points	CLPM	factor CLPM	RI-CLPM	STARTS	LCM-SR	ALT	LCS
stationarity	T=2	14	17	17	20	24	34	24
	T=3	16	19	19	22	24	34	24
	T=4	18	21	21	24	24	34	24
	T=5	20	23	23	26	24	34	24
non-stationarity	T=2	14	17	17	20	24	34	24
	T=3	23	26	26	29	31	41	31
	T=4	32	35	35	38	38	48	38
	T=5	41	44	44	47	45	55	45
Minimum T in stationarity assumption		2	3	3	3	3	4	3
Minimum T in non-stationarity assumption		2	3	3	4	4	5	4

*The numbers of second-order moments are 14, 27, 44 and 65 for T=2,3,4 and 5, respectively.

*The assumption of stability indicates that autoregressive parameters, cross-lagged (or, coupling) parameters, residual and error variances and covariances are time-invariant, while instability indicates time-variant.

*In the LCM-SR, factor loadings from slope factors S are assumed to be prespecified as αt= t-1.

*In the ALT model, factor loadings from accumulating factors B are assumed to be prespecified as αt= t-1.

*For non-stationarity condition, stationarity constraint of unique factor (co)variances is imposed in the factor CLPM, STARTS and the LCS models for identification.

Table 2. An overview of cross-lagged models that indicates which of components they include.

	Measurement equations	Decomposition equations			Dynamic equations		
Model	ϵ	μ	I	S	A	B	d
CLPM	$-$	$+$	$-$	$-$	$-$	$-$	$+$
factor CLPM	$+$	$+$	$-$	$-$	$-$	$-$	$+$
RI-CLPM	$-$	$+$	$+^{(a)}$	$-$	$-$	$-$	$+$
STARTS	$+$	$+$	$+^{(a)}$	$-$	$-$	$-$	$+$
LCM-SR	$-$	$-$	$+$	$+$	$-$	$-$	$+$
ALT	$-$	$-$	$-$	$-$	$+$	$+$	$+$
LCS	$+$	$-$	$-$	$-$	$+$	$-$	$+^{(b)}$

$+^{(a)}$ These factors have a mean of zero, unless the μ' s are fixed to zero.

$+^{(b)}$ This residual is typically fixed to zero for estimation purposes.

Table 3. Overview of the bivariate longitudinal models considered in the current study (for reasons of space, we only include the expression for one of the two variables).

Acronym	Equations	Purpose
Unified framework	$x_{it} = f_{xit} + \epsilon_{xit}$	Measurement equation
	$f_{xit} = [\mu_{xt} + \{I_{xi} + (t-1)S_{xi}\}] + f^*_{xit}$	Decompostion equation
	$f^*_{xit} = \{A_{xi} + (t-1)B_{xi}\} + \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Dynamic equation
CLPM	$x_{it} = f_{xit}$	Identical
	$f_{xit} = \mu_{xt} + f^*_{xit}$	Dissociation of temporal means
	$f^*_{xit} = \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression
factor CLPM	$x_{it} = f_{xit} + \epsilon_{xit}$	Measurement errors
	$f_{xit} = \mu_{xt} + f^*_{xit}$	Dissociation of temporal means
	$f^*_{xit} = \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression
RI-CLPM	$x_{it} = f_{xit}$	Identical
	$f_{xit} = [\mu_{xt} + I_{xi}] + f^*_{xit}$	Dissociation of temporal and person means
	$f^*_{xit} = \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression
STARTS model	$x_{it} = f_{xit} + \epsilon_{xit}$	Measurement errors
	$f_{xit} = [\mu_{xt} + I_{xi}] + f^*_{xit}$	Dissociation of temporal and person means
	$f^*_{xit} = \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression
LCM-SR	$x_{it} = f_{xit}$	Identical
	$f_{xit} = [I_{xi} + (t-1)S_{xi}] + f^*_{xit}$	Growth curves
	$f^*_{xit} = \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression
ALT model	$x_{it} = f_{xit}$	Identical
	$f_{xit} = f^*_{xit}$	No detrending
	$f^*_{xit} = \{A_{xi} + (t-1)B_{xi}\} + \beta_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression and growth curves
LCS model	$x_{it} = f_{xit} + \epsilon_{xit}$	Measurement errors
	$f_{xit} = f^*_{xit}$	No detrending
	$f^*_{xit} = A_{xi} + \beta^*_x f^*_{xi(t-1)} + \gamma_x f^*_{yi(t-1)} + d_{xit}$	Lagged regression and growth curves

*We supposed time-invariant autoregressive and cross-lagged parameters in all the models here.

*$\beta^*_x = 1 + \beta_x$.

Table 4. Parameter estimates and model fit indices from different cross-lagged models (N=4,670)

	CLPM			factor CLPM			RI-CLPM			LCM-SR			LCS		
	Est.	*SE*	*p*	Est.	*SE*	*p*	Est.	*SE*	*p*	Est.	*SE*	*p*	Est.	*SE*	*p*
βy	0.818	0.006	0.000	0.955	0.010	0.000	0.700	0.012	0.000	0.349	0.027	0.000	0.792	0.033	0.000
γy	0.037	0.013	0.003	0.042	0.019	0.024	-0.005	0.022	0.824	0.026	0.025	0.308	-0.131	0.094	0.163
βx	0.511	0.006	0.000	0.861	0.008	0.000	0.166	0.011	0.000	0.082	0.014	0.000	0.736	0.047	0.000
γx	0.009	0.004	0.022	-0.004	0.004	0.349	-0.008	0.007	0.250	0.015	0.011	0.178	-0.019	0.009	0.042
CFI	0.875			0.970			0.959			0.973			0.973		
TLI	0.879			0.970			0.958			0.973			0.973		
AIC	73465.636			72051.655			72224.166			72009.638			72015.863		
BIC	73607.512			72212.878			72385.389			72164.412			72170.637		
RMSEA [95% CI]	0.077 [0.074, 0.079]			0.038 [0.035, 0.041]			0.045 [0.042, 0.048]			0.036 [0.033, 0.039]			0.036 [0.033, 0.039]		
SRMR	0.113			0.047			0.066			0.042			0.044		
degree of freedom	68			65			65			66			66		
Num of Parameters	22			25			25			24			24		

*CFI…comparative fit index, TLI…Tucker-Levis index, AIC…Akaike information criterion, BIC…Bayesian information criterion, RMSEA…root mean square error of approximation, SRMR…standardized root mean square residual.

*The STARTS model and the ALT model both resulted in an improper solution.

Figure Captions

Figure 1. Path diagrams of cross-lagged models.

Notes. Residuals covariances and covariances between common factors are all omitted for clarity of presentation. Note that means of common (trait) factors are set to zero in the RI-CLPM and the STARTS model.

Figure 2. A conceptual diagram to clarify the relations among cross-lagged models.

Notes. One-headed arrow indicates nested relations, while dotted line indicates that relationship can be satisfied conditionally. Double-headed dotted line indicates that models are statistically equivalent under particular circumstances. Note that we supposed time-invariant autoregressive and cross-lagged parameters in all the models here.

Notes. CLPM…cross-lagged panel model, RI-CLPM…random-intercepts CLPM, STARTS…stable trait autoregressive trait and state, LCM-SR…latent curve model with structured residuals, ALT…autoregressive latent trajectory, LCS…latent change score, TCS…triple change score.

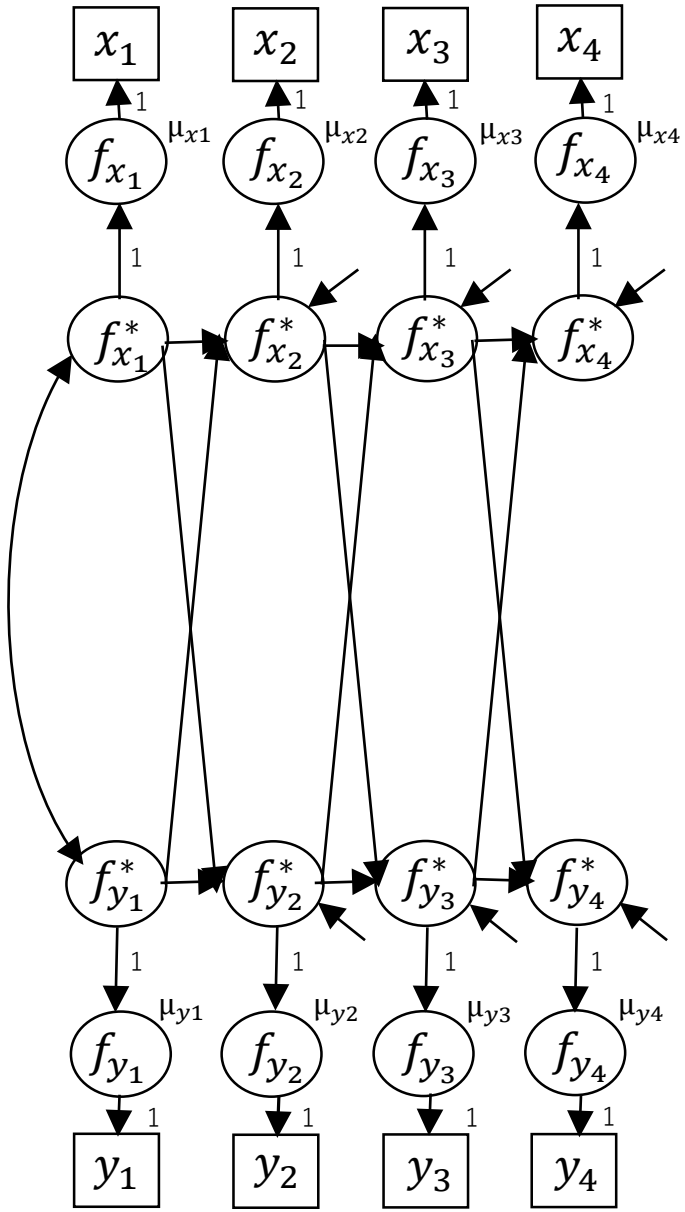

Figure 1a. CLPM

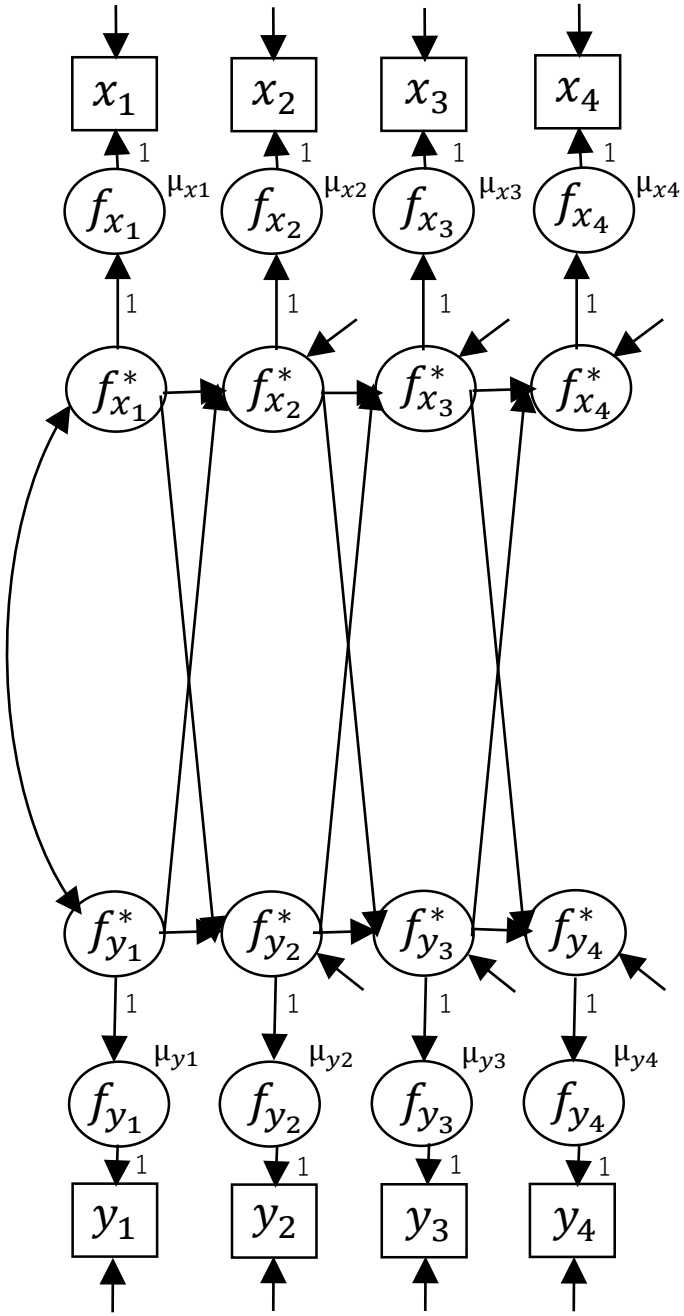

Figure 1b. factor CLPM

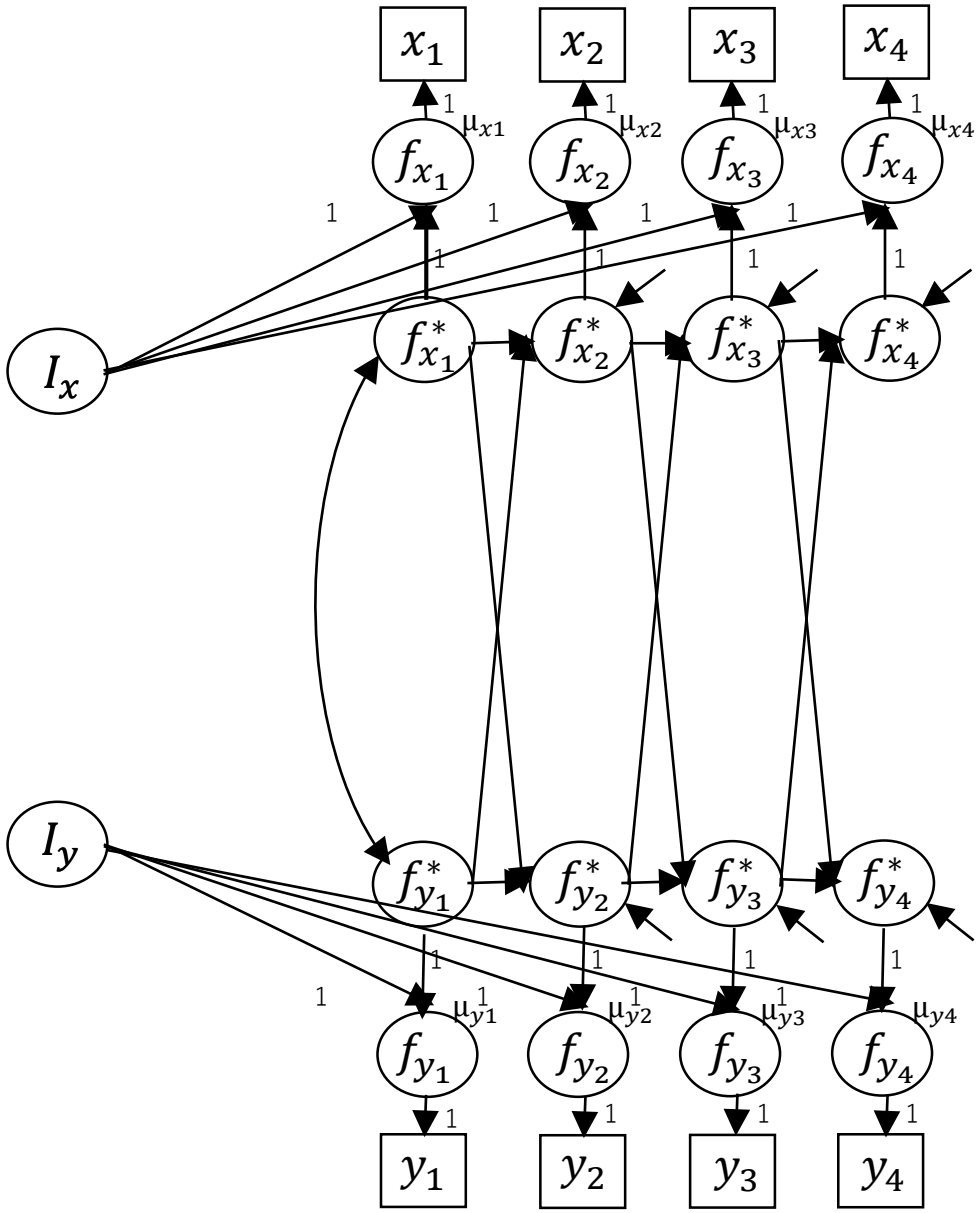

Figure 1c. RI-CLPM

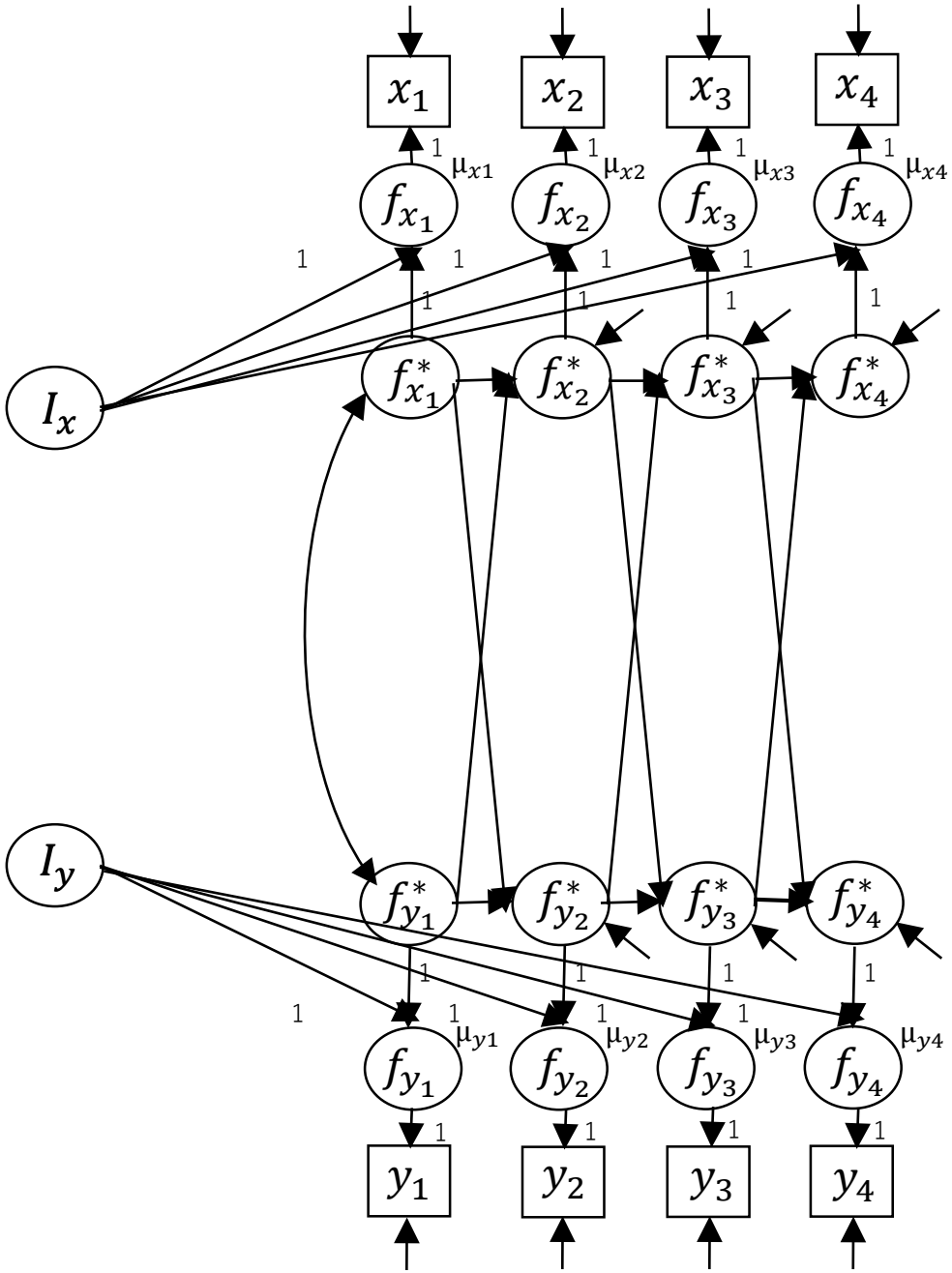

Figure 1d. STARTS model

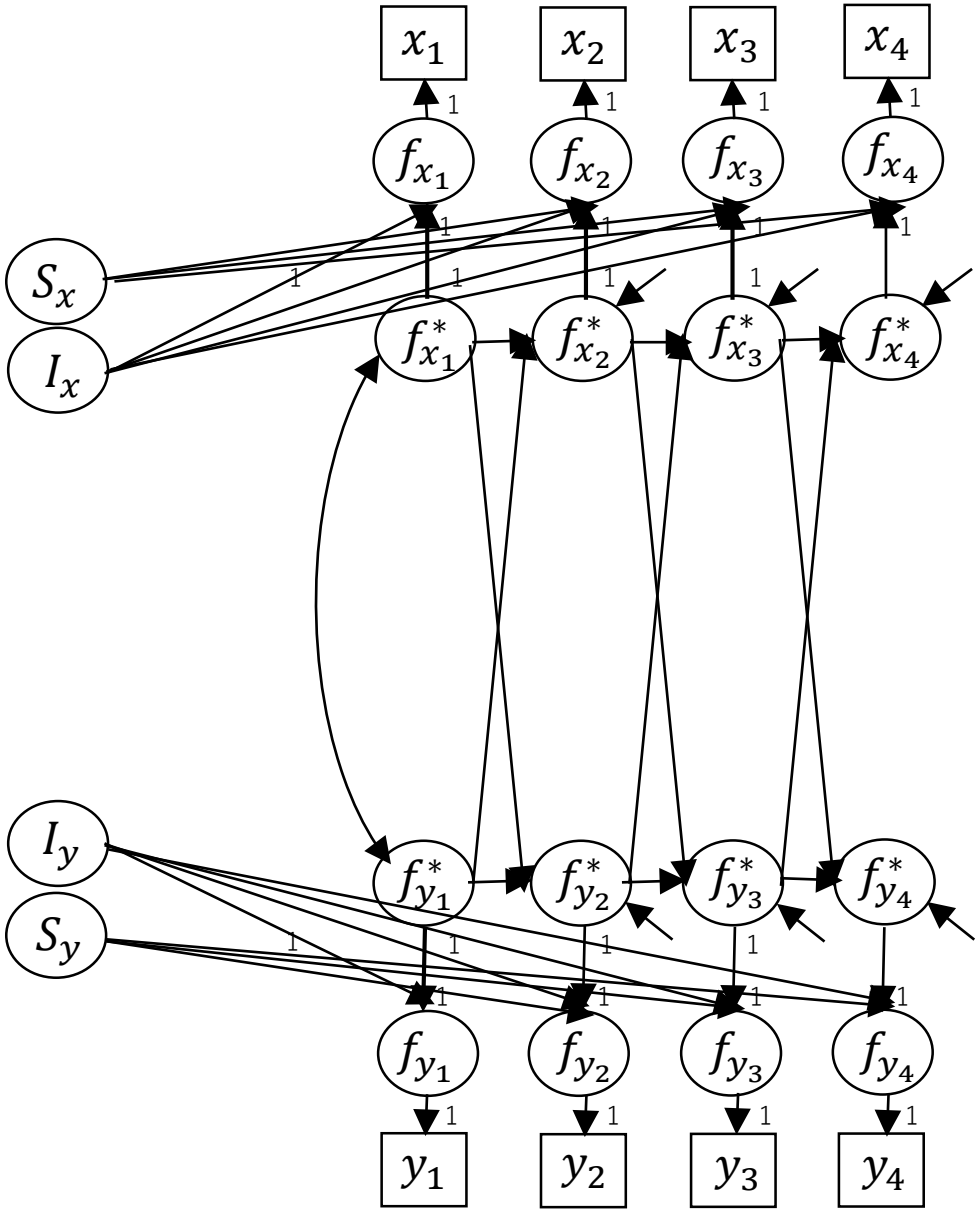

Figure 1e. LCM-SR

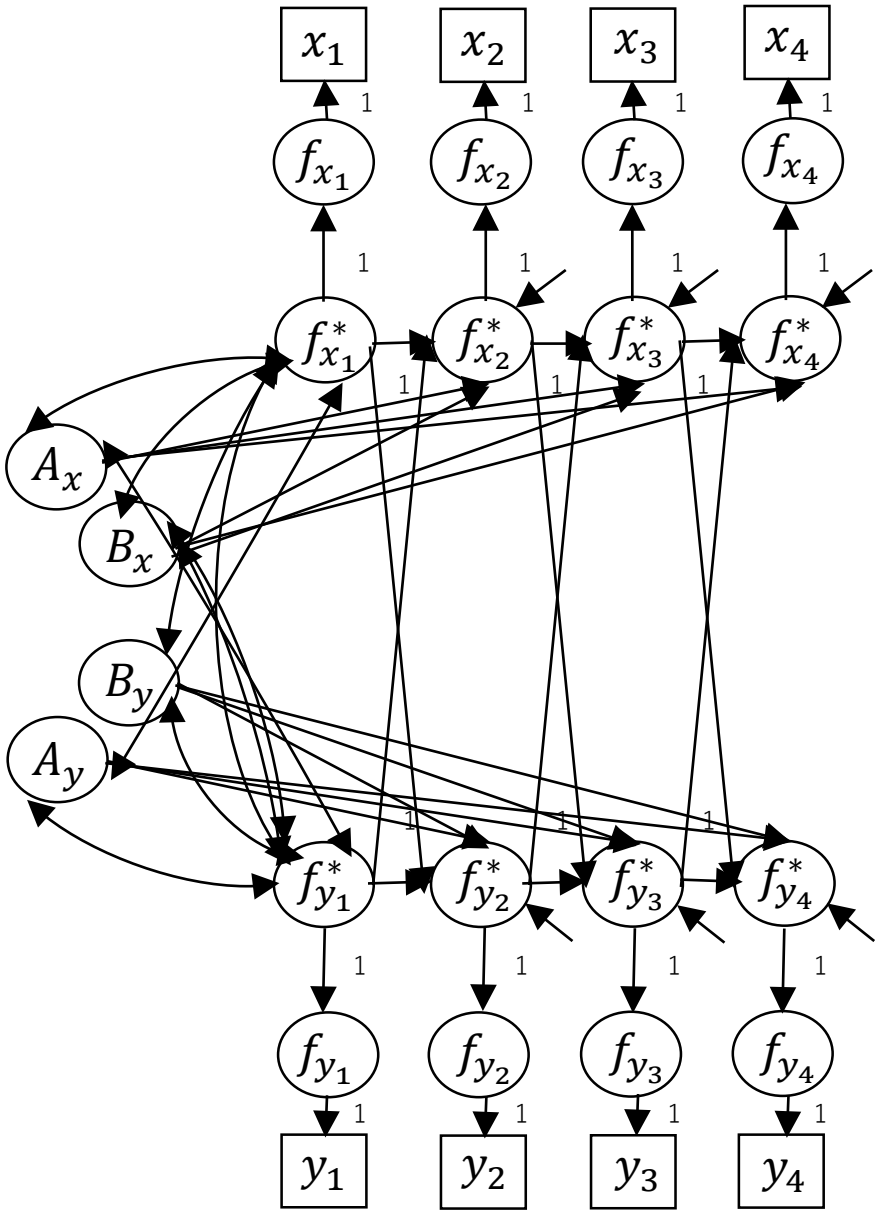

Figure 1f. ALT model

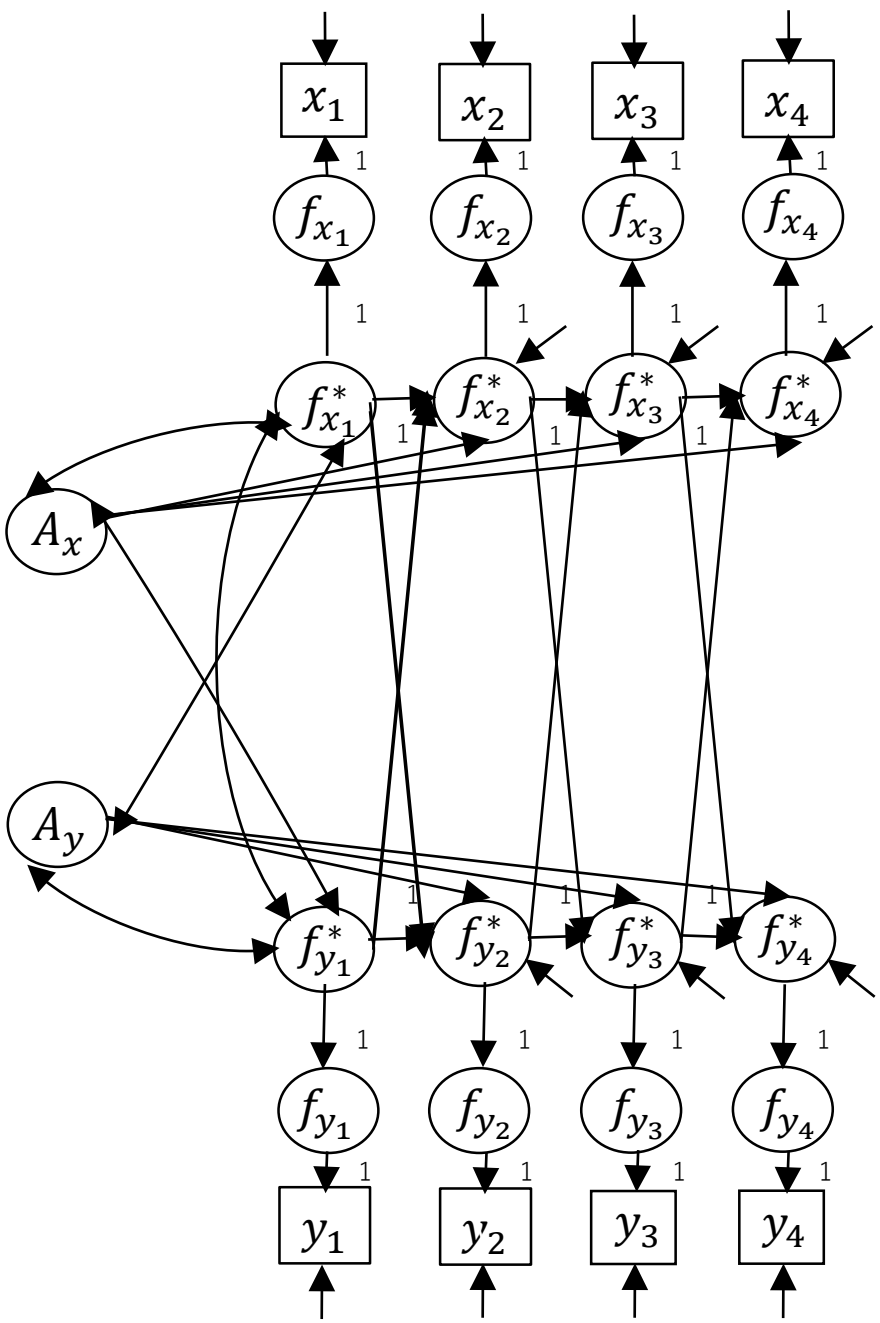

Figure 1g. LCS model

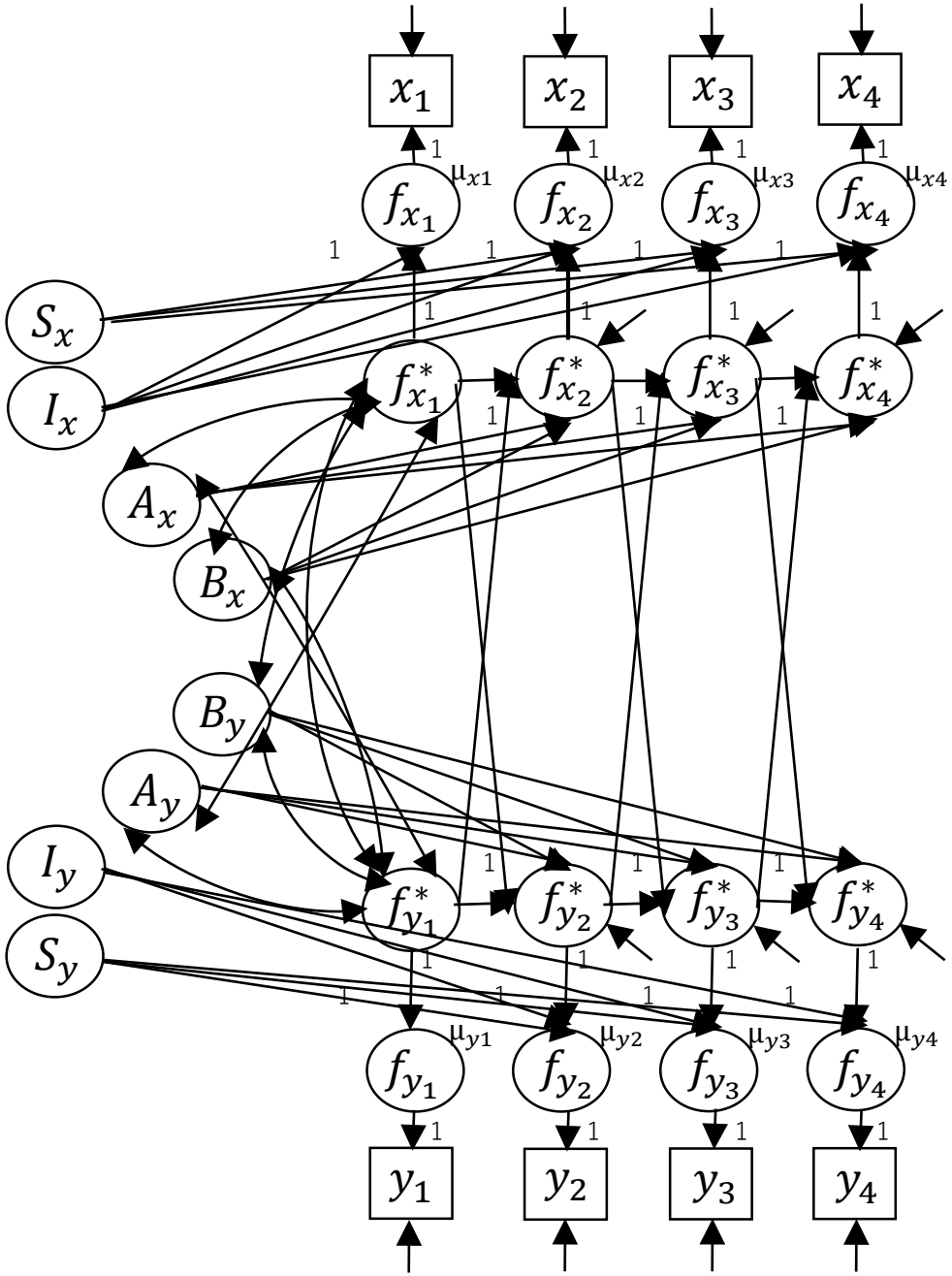

Figure 1h. Unified framework

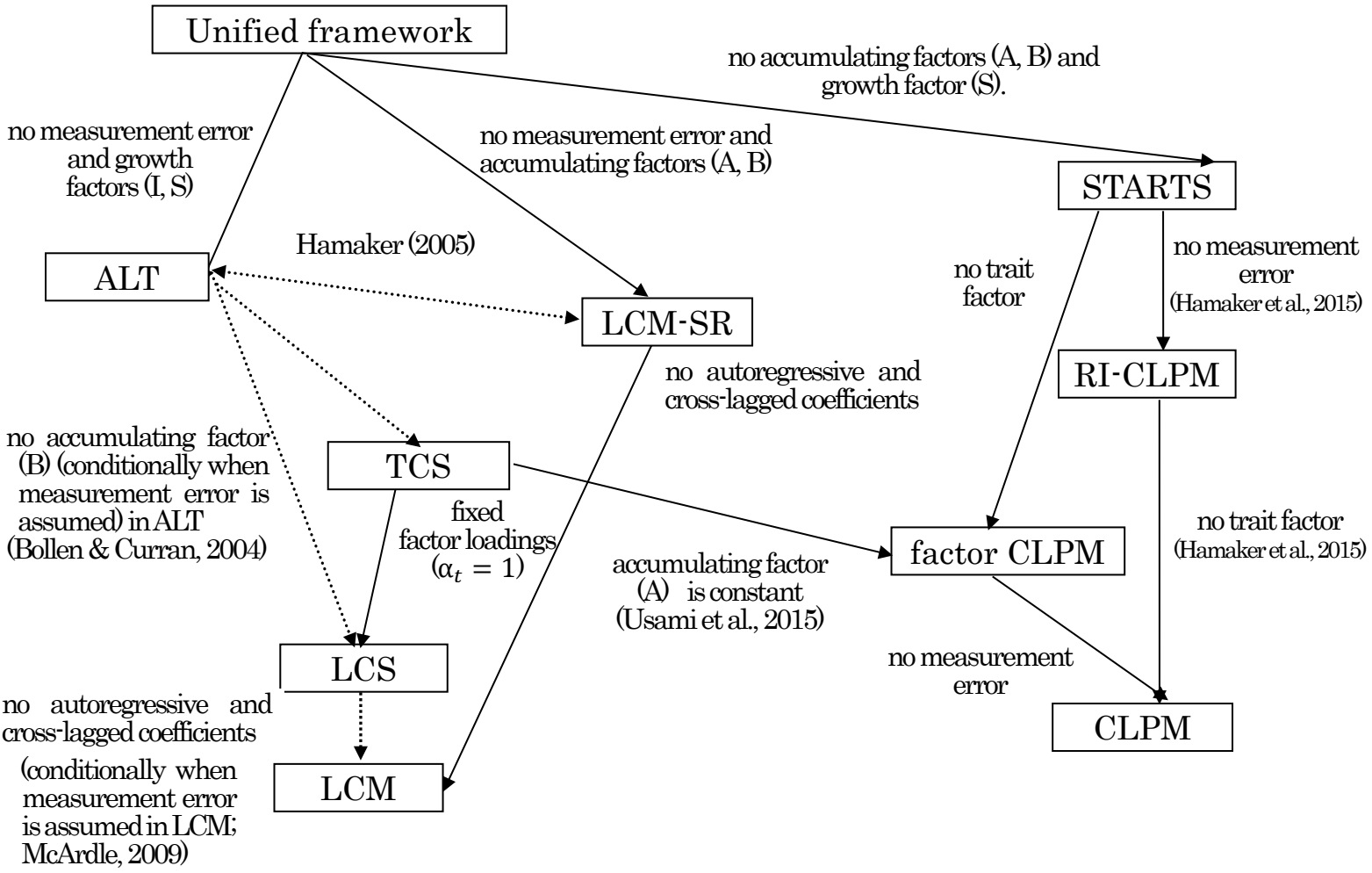
Unified framework
no accumulating factors (A, B) and growth factor (S).
no measurement error and growth factors (I, S)
no measurement error and accumulating factors (A, B)
STARTS
Hamaker (2005)
ALT
no trait factor
no measurement error (Hamaker et al., 2015)
LCM-SR
RI-CLPM
no autoregressive and cross-lagged coefficients
no accumulating factor (B) (conditionally when measurement error is assumed) in ALT (Bollen & Curran, 2004)
TCS
fixed factor loadings ($\alpha_t = 1$)
factor CLPM
no trait factor (Hamaker et al., 2015)
accumulating factor (A) is constant (Usami et al., 2015)
LCS
no measurement error
CLPM
no autoregressive and cross-lagged coefficients (conditionally when measurement error is assumed in LCM; McArdle, 2009)
LCM

Figure 2

Online Supplemental Material

Sample means and covariance matrix in empirical data example… p.2

The lavaan source codes…… p.3

Online Supplemental Material: Sample means and covariance matrix in empirical data example

Means

X: perceived exposure to smoking in movies						Y: smoking intensity					
Age 15	Age 16	Age 17	Age 18	Age 19	Age 20	Age 15	Age 16	Age 17	Age 18	Age 19	Age 20
3.140	3.140	3.078	3.068	3.046	3.029	1.219	1.346	1.463	1.692	1.794	1.867

Covariance matrix

		X: perceived exposure to smoking in movies						Y: smoking intensity					
		Age 15	Age 16	Age 17	Age 18	Age 19	Age 20	Age 15	Age 16	Age 17	Age 18	Age 19	Age 20
X: perceived exposure to smoking in movies	Age 15	0.414	0.207	0.173	0.159	0.147	0.127	0.081	0.077	0.076	0.086	0.087	0.096
	Age 16	0.207	0.392	0.185	0.169	0.147	0.144	0.112	0.080	0.064	0.085	0.089	0.094
	Age 17	0.173	0.185	0.349	0.182	0.152	0.154	0.084	0.079	0.080	0.087	0.078	0.067
	Age 18	0.159	0.169	0.182	0.341	0.170	0.168	0.033	0.037	0.034	0.074	0.046	0.044
	Age 19	0.147	0.147	0.152	0.170	0.324	0.181	0.247	0.108	0.052	0.066	0.070	0.067
	Age 20	0.127	0.144	0.154	0.168	0.181	0.327	NA	0.231	0.037	0.063	0.041	0.061
Y: smoking intensity	Age 15	0.081	0.112	0.084	0.033	0.247	NA	0.649	0.576	0.525	0.572	0.461	0.719
	Age 16	0.077	0.080	0.079	0.037	0.108	0.231	0.576	0.974	0.723	0.665	0.599	0.677
	Age 17	0.076	0.064	0.080	0.034	0.052	0.037	0.525	0.723	1.181	0.923	0.851	0.846
	Age 18	0.086	0.085	0.087	0.074	0.066	0.063	0.572	0.665	0.923	1.720	1.322	1.230
	Age 19	0.087	0.089	0.078	0.046	0.070	0.041	0.461	0.599	0.851	1.322	1.923	1.508
	Age 20	0.096	0.094	0.067	0.044	0.067	0.061	0.719	0.677	0.846	1.230	1.508	2.064

*NA⋯not available

Online Supplemental Material: The lavaan source codes

***Original data (denoted as "dat" in the code below) can be available through the website of Inter-university Consortium for Political and Social Research (https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36282).**

*** The lavaan package has to be installed in advance. To install lavaan, simply start up R, and type:**

```
install.packages("lavaan", dependencies = TRUE)
```

Data cleaning and retrieving

```
attach(dat)
dat2<-subset(dat, AGE_YRS>14 & AGE_YRS<21)    #age restriction
detach(dat); attach(dat2)
Data<-data.frame(matrix(rep(NA,13*length(table(MASTERID))),length(table(MASTERID)),13))
KK<- MASTERID[1]; AA<-1
 for(j in 1:nrow(dat2)){    #Creating data as subject×time format
 KK2<- MASTERID[j]
 AA<-ifelse(KK2==KK,AA,AA+1)
 Data[AA,1]<-MASTERID[j]    #ID
 Data[AA,AGE_YRS[j]-13]<-Q0680[j]    # perceived exposure to smoking in movies
 Data[AA,AGE_YRS[j]-7]<-SMOKEINTENSITY[j]    # smoking intensity
 KK<-KK2
 }
detach(dat2)

colnames(Data)<-c("ID","x1","x2","x3","x4","x5","x6" ,"y1","y2","y3","y4","y5","y6")
attach(Data)
Data<-data.frame(cbind(x1,x2,x3,x4,x5,x6,y1,y2,y3,y4,y5,y6))
```

#Lavaan source codes for each cross-lagged model (T=6).

```
#CLPM(T=6)
 CLPMT6 <- '
      x1~mux1*1;  y1~muy1*1;  x2~mux2*1;  y2~muy2*1;  x3~mux3*1;  y3~muy3*1;  x4~mux4*1;
      y4~muy4*1;   x5~mux5*1; y5~muy5*1; x6~mux6*1; y6~muy6*1
      FFx1~0*1;FFy1~0*1; FFx1~~phix*FFx1; FFy1~~phiy*FFy1; FFx1~~phixy*FFy1;
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;  FFx3~~Omegax*FFx3;  FFx4~~Omegax*FFx4;  FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6
      FFy2~~Omegay*FFy2;  FFy3~~Omegay*FFy3;  FFy4~~Omegay*FFy4;  FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
'
```

```
#factor-CLPM(T=6)
 factorCLPMT6 <- '
      x1~mux1*1;  y1~muy1*1;  x2~mux2*1;  y2~muy2*1;  x3~mux3*1;  y3~muy3*1;  x4~mux4*1;
      y4~muy4*1;   x5~mux5*1; y5~muy5*1; x6~mux6*1; y6~muy6*1
      x1 ~~ psix*x1; x2 ~~ psix*x2; x3 ~~ psix*x3; x4 ~~ psix*x4; x5 ~~ psix*x5; x6 ~~ psix*x6
      y1 ~~ psiy*y1; y2 ~~ psiy*y2; y3 ~~ psiy*y3; y4 ~~ psiy*y4; y5 ~~ psiy*y5; y6 ~~ psiy*y6
      y1 ~~ psixy*x1; y2 ~~ psixy*x2; y3 ~~ psixy*x3; y4 ~~ psixy*x4; y5 ~~ psixy*x5; y6 ~~ psixy*x6
      FFx1~0*1;FFy1~0*1; FFx1~~phix*FFx1; FFy1~~phiy*FFy1; FFx1~~phixy*FFy1;
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;  FFx3~~Omegax*FFx3;  FFx4~~Omegax*FFx4;  FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6
      FFy2~~Omegay*FFy2;  FFy3~~Omegay*FFy3;  FFy4~~Omegay*FFy4;  FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
'

#RI-CLPM(T=6)
 RICLPMT6 <- '
      x1~mux1*1;  y1~muy1*1;  x2~mux2*1;  y2~muy2*1;  x3~mux3*1;  y3~muy3*1;  x4~mux4*1;
      y4~muy4*1;   x5~mux5*1; y5~muy5*1; x6~mux6*1; y6~muy6*1
      x1~~0*x1;  y1~~0*y1;  x2~~0*x2;  y2~~0*y2;  x3~~0*x3;  y3~~0*y3;  x4~~0*x4;  y4~~0*y4;
      x5~~0*x5; y5~~0*y5; x6~~0*x6; y6~~0*y6;
      Tx~~0*FFx1;Tx~~0*FFy1; Ty~~0*FFx1;Ty~~0*FFy1;
      Tx=~1*x1+1*x2+1*x3+1*x4+1*x5+1*x6; Ty=~1*y1+1*y2+1*y3+1*y4+1*y5+1*y6
      Tx~0*1; Ty~0*1; Tx~~taux*Tx; Ty~~tauy*Ty; Tx~~tauxy*Ty
      FFx1~0*1;FFy1~0*1; FFx1~~phix*FFx1; FFy1~~phiy*FFy1; FFx1~~phixy*FFy1;
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;  FFx3~~Omegax*FFx3;  FFx4~~Omegax*FFx4;  FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6
      FFy2~~Omegay*FFy2;  FFy3~~Omegay*FFy3;  FFy4~~Omegay*FFy4;  FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
'

#STARTS(T=6)
 STARTST6 <- '
      x1~mux1*1;  y1~muy1*1;  x2~mux2*1;  y2~muy2*1;  x3~mux3*1;  y3~muy3*1;  x4~mux4*1;
      y4~muy4*1;   x5~mux5*1; y5~muy5*1; x6~mux6*1; y6~muy6*1
      Tx=~1*x1+1*x2+1*x3+1*x4+1*x5+1*x6; Ty=~1*y1+1*y2+1*y3+1*y4+1*y5+1*y6
      Tx~~0*FFx1;Tx~~0*FFy1; Ty~~0*FFx1;Ty~~0*FFy1;
      Tx~0*1; Ty~0*1; Tx~~taux*Tx; Ty~~tauy*Ty; Tx~~tauxy*Ty
      x1 ~~ psix*x1; x2 ~~ psix*x2; x3 ~~ psix*x3; x4 ~~ psix*x4; x5 ~~ psix*x5; x6 ~~ psix*x6
      y1 ~~ psiy*y1; y2 ~~ psiy*y2; y3 ~~ psiy*y3; y4 ~~ psiy*y4; y5 ~~ psiy*y5; y6 ~~ psiy*y6
```

```
      y1 ~~ psixy*x1; y2 ~~ psixy*x2; y3 ~~ psixy*x3; y4 ~~ psixy*x4; y5 ~~ psixy*x5; y6 ~~ psixy*x6
      FFx1~0*1;FFy1~0*1; FFx1~~phix*FFx1; FFy1~~phiy*FFy1; FFx1~~phixy*FFy1;
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;  FFx3~~Omegax*FFx3;  FFx4~~Omegax*FFx4;  FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6
      FFy2~~Omegay*FFy2;  FFy3~~Omegay*FFy3;  FFy4~~Omegay*FFy4;  FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
'

#LCM-SR(T=6)
 LCMSRT6 <- '
      Ix=~1*x1+1*x2+1*x3+1*x4+1*x5+1*x6; Iy=~1*y1+1*y2+1*y3+1*y4+1*y5+1*y6
      Sx=~0*x1+1*x2+2*x3+3*x4+4*x5+5*x6; Sy=~0*y1+1*y2+2*y3+3*y4+4*y5+5*y6
      x1~0*1; y1~0*1; x2~0*1; y2~0*1; x3~0*1; y3~0*1; x4~0*1; y4~0*1;x5~0*1; y5~0*1; x6~0*1;
      y6~0*1
      x1~~0*x1; y1~~0*y1; x2~~0*x2; y2~~0*y2; x3~~0*x3; y3~~0*y3; x4~~0*x4; y4~~0*y4;
      x5~~0*x5; y5~~0*y5; x6~~0*x6; y6~~0*y6;
      Ix~muIx*1; Iy~muIy*1; Sx~muSx*1; Sy~muSy*1
      FFx1~0*1; FFx1~~0*Ix;FFx1~~0*Iy;FFx1~~0*Sx;FFx1~~0*Sy
      FFy1~0*1; FFy1~~0*Ix;FFy1~~0*Iy;FFy1~~0*Sx;FFy1~~0*Sy
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;  FFx3~~Omegax*FFx3;  FFx4~~Omegax*FFx4;  FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6
      FFy2~~Omegay*FFy2;  FFy3~~Omegay*FFy3;  FFy4~~Omegay*FFy4;  FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
'


#ALT(T=6)
 ALTT6 <- '
      x1~0*1; y1~0*1; x2~0*1; y2~0*1; x3~0*1; y3~0*1; x4~0*1; y4~0*1;x5~0*1; y5~0*1; x6~0*1;
      y6~0*1
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
      Iy=~1*FFy2+1*FFy3+1*FFy4+1*FFy5+1*FFy6; Ix=~1*FFx2+1*FFx3+1*FFx4+1*FFx5+1*FFx6
      Sy=~1*FFy2+2*FFy3+3*FFy4+4*FFy5+5*FFy6; Sx=~1*FFx2+2*FFx3+3*FFx4+4*FFx5+5*FFx6
      FFx1~mux1*1; FFy1~muy1*1; Ix~muIx*1;Iy~muIy*1; Sx~muSx*1;Sy~muSy*1
      FFx1~~phix1*  FFx1;  FFy1~~phiy1*FFy1;  Ix~~phiIx*Ix;Iy~~phiIy*Iy;
      Sx~~phiSx*Sx;Sy~~phiSy*Sy
      FFx1~~phix1Sx*Sx; FFx1~~phix1Sy*Sy;
      FFy1~~phiy1Sx*Sx; FFy1~~phiy1Sy*Sy;
```

```
      Ix~~phiIxSx*Sx; Ix~~phiIxSy*Sy;
      Iy~~phiIySx*Sx; Iy~~phiIySy*Sy; Sx~~phiSxSy*Sy
      FFx1~~phix1y1* FFy1; FFx1~~phix1Ix*Ix; FFx1~~phix1Iy*Iy;FFy1~~phiy1Ix*Ix
      FFy1~~phiy1Iy*Iy; Ix~~phiIxIy*Iy
      x1 ~~ 0*x1; x2 ~~ 0*x2; x3 ~~ 0*x3; x4 ~~ 0*x4; x5 ~~ 0*x5; x6 ~~ 0*x6
      y1 ~~ 0*y1; y2 ~~ 0*y2; y3 ~~ 0*y3; y4 ~~ 0*y4; y5 ~~ 0*y5; y6 ~~ 0*y6
      y1 ~~ 0*x1; y2 ~~ 0*x2; y3 ~~ 0*x3; y4 ~~ 0*x4; y5 ~~ 0*x5; y6 ~~ 0*x6
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;   FFx3~~Omegax*FFx3;   FFx4~~Omegax*FFx4;   FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6;
      FFy2~~Omegay*FFy2;   FFy3~~Omegay*FFy3;   FFy4~~Omegay*FFy4;   FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
'


#LCS(T=6)
 LCST6 <- '
      x1~0*1; y1~0*1; x2~0*1; y2~0*1; x3~0*1; y3~0*1; x4~0*1; y4~0*1;x5~0*1; y5~0*1; x6~0*1;
      y6~0*1
      FFy1 =~ 1*y1; FFy2 =~ 1*y2; FFy3 =~ 1*y3; FFy4 =~ 1*y4; FFy5 =~ 1*y5; FFy6 =~ 1*y6
      FFx1 =~ 1*x1; FFx2 =~ 1*x2; FFx3 =~ 1*x3; FFx4 =~ 1*x4; FFx5 =~ 1*x5; FFx6 =~ 1*x6
      Iy=~1*FFy2+1*FFy3+1*FFy4+1*FFy5+1*FFy6; Ix=~1*FFx2+1*FFx3+1*FFx4+1*FFx5+1*FFx6
      FFx1~mux1*1; FFy1~muy1*1; Ix~muIx*1;Iy~muIy*1
      FFx1~~phix1* FFx1; FFy1~~phiy1*FFy1; Ix~~phiIx*Ix;Iy~~phiIy*Iy
      FFx1~~phix1y1* FFy1; FFx1~~phix1Ix*Ix; FFx1~~phix1Iy*Iy;FFy1~~phiy1Ix*Ix
      FFy1~~phiy1Iy*Iy; Ix~~phiIxIy*Iy
      x1 ~~ psix*x1; x2 ~~ psix*x2; x3 ~~ psix*x3; x4 ~~ psix*x4; x5 ~~ psix*x5; x6 ~~ psix*x6
      y1 ~~ psiy*y1; y2 ~~ psiy*y2; y3 ~~ psiy*y3; y4 ~~ psiy*y4; y5 ~~ psiy*y5; y6 ~~ psiy*y6
      y1 ~~ psixy*x1; y2 ~~ psixy*x2; y3 ~~ psixy*x3; y4 ~~ psixy*x4; y5 ~~ psixy*x5; y6 ~~ psixy*x6
      FFy2~ betay*FFy1+gammay*FFx1; FFx2~ betax*FFx1+gammax*FFy1
      FFy3~ betay*FFy2+gammay*FFx2; FFx3~ betax*FFx2+gammax*FFy2
      FFy4~ betay*FFy3+gammay*FFx3; FFx4~ betax*FFx3+gammax*FFy3
      FFy5~ betay*FFy4+gammay*FFx4; FFx5~ betax*FFx4+gammax*FFy4
      FFy6~ betay*FFy5+gammay*FFx5; FFx6~ betax*FFx5+gammax*FFy5
      FFx2~~Omegax*FFx2;   FFx3~~Omegax*FFx3;   FFx4~~Omegax*FFx4;   FFx5~~Omegax*FFx5;
      FFx6~~Omegax*FFx6;
      FFy2~~Omegay*FFy2;   FFy3~~Omegay*FFy3;   FFy4~~Omegay*FFy4;   FFy5~~Omegay*FFy5;
      FFy6~~Omegay*FFy6
      FFx2~~Omegaxy*FFy2; FFx3~~Omegaxy*FFy3; FFx4~~Omegaxy*FFy4; FFx5~~Omegaxy*FFy5;
      FFx6~~Omegaxy*FFy6
'
```

```
library(MASS); require("lavaan")

fit<-cfa(CLPMT6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(factorCLPMT6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(RICLPMT6,data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(STARTST6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(LCMSRT6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(ALTT6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
fit<-cfa(LCST6, data=Data, missing='fiml')
summary(fit, fit.measures = TRUE)
```