

## **A COMPARISON OF THREE APPROACHES FOR EVALUATING AUTOREGRESSIVE AND CROSS LAGGED EFFECTS**

**YI PAN<sup>1</sup>, QIU WANG<sup>2</sup> and MICHAEL THOMAS WILLOUGHBY<sup>3</sup>**

<sup>1</sup>Frank Porter Graham Child Development Institute  
University of North Carolina at Chapel Hill  
USA

<sup>2</sup>School of Education  
Syracuse University  
USA  
e-mail: wangqiu@syr.edu

<sup>3</sup>RTI International  
3040 East Cornwallis Road  
Post Office Box 12194  
Research Triangle Park  
NC 27709-2194  
USA

### **Abstract**

Researchers in the social and behavioural sciences often ask questions regarding the time-linked associations between two or more constructs at multiple assessments across time. Although their questions typically involve an interest in making within-subject inferences, the most commonly used analytic approach does not disaggregate between-from within-subject variations, resulting in a mismatch between question and method. In this paper, we compare and

---

2010 Mathematics Subject Classification: 62P15, 62-07.

Keywords and phrases: autoregressive model, autoregressive latent trajectory model, cross-lagged model, fixed influence cross-lagged model, repeated measures analysis, structural equation model.

Received March 28, 2015

contrast three analytic approaches that can be used to test time-linked (typically cross-lagged) associations between two variables that are repeatedly measured over time. Through use of a real-world example and Monte Carlo simulation, we demonstrate how two of these approaches facilitate within-subject inferences. We also provide practical advice to researchers to help determine which analytic approach is most appropriate for their research question.

## 1. Introduction

Researchers in the social, behaviour, educational and health sciences frequently ask questions regarding whether and how two or more constructs are related to each other across time. Typically, questions of sort emerge from studies employing passive longitudinal designs, in which two or more constructs of interest are measured at multiple assessments. In some situations, the substantive question involves whether and how these two constructs “travel together” across time. That is, the focus is on whether the rate of change in one domain is related to the rate of change in another domain. Multivariate growth curve modelling approaches provide one means for answering questions of this nature. There are a variety of pedagogical papers that have elaborated these sorts of models (Ferrer & McArdle [10]; Grimm [12]; MacCallum et al. [17]; Sayer & Willett [24]).

A more common situation involves interest in understanding the time-linked associations between two or more constructs at multiple assessments. Whereas the former question typically uses slope (or difference) terms to summarize changes in two or more domains, the latter question typically focuses on the cross lagged associations between two constructs across assessments (e.g.,  $X_1 \rightarrow Y_2$ ,  $Y_1 \rightarrow X_2$ ). For much of the past 25 years, autoregressive cross lagged (ARCL) models have been the dominant strategy for approaching questions of this latter type. A number of didactic presentations of ARCL models are available (Farrell [9]; Mayer & Carroll [19]).

Even a quick perusal of the recent behavioural, educational and health science literatures reveals that ARCL models are in wide use. For example, behavioural researchers have used ARCL models to test the reciprocal associations between depressive symptomatology and friendship networks (Chan & Poulin [6]), educational researchers have studied the reciprocal associations between student beliefs and student learning (Phan [21]), substance use researchers have investigated the reciprocal associations between perceived social norms and drinking behaviours (Cullum et al. [7]), health researchers have tested the reciprocal associations between psychological coping strategies and biological indices of glycemic control (Luyckx et al. [16]), and personality psychologists have tested the reciprocal associations between cognitive and belief orientations as they relate to prejudice and discrimination behaviours (Kteily et al. [15]). A ubiquitous feature of studies employing ARCL models is an interest in understanding the cross lagged associations between two constructs controlling for earlier levels of each construct, and whether one construct is “privileged” in its association with the other. That is, to the extent that one construct is uniquely predictive of the other but the converse is not true (i.e.,  $X_1 \rightarrow Y_2$  is statistically significant but  $Y_1 \rightarrow X_2$  is not), researchers often infer that one construct is causally related to the other, even though strong causal language is rarely used.

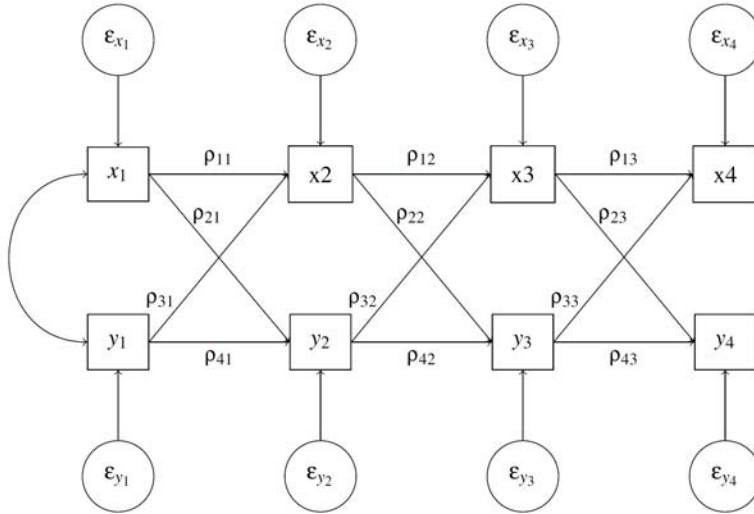
It is our position that in many cases, the inferences that social and behavioural researchers draw from their ARCL models are troubled with potential estimation biases. This derives from the fact that although researchers are typically interested in making within-subject inferences about the constructs under study (e.g., changing perceptions of drinking climate will result in lower levels of drinking; changing one’s belief orientations will reduce prejudiced behaviours), the standard application of ARCL models does not appropriately parse between-and within-subject variations. Moreover, ARCL models have difficulties in controlling for unmeasured variables that may confound any observed associations between the two constructs under study as are often assumed. In this paper, we describe and contrast three analytic approaches, the

autoregressive cross lagged (ARCL) model (Mayer & Carroll [19]), the autoregressive latent trajectory (ALT) model (Bollen & Curran [4]), and the fixed influence cross lagged (FICL) model (Allison [2]), which permit researchers to answer questions about the time-linked associations between two (or more) constructs measured at multiple assessments. We apply all three approaches to a real world dataset and point out discrepancies in conclusions that result from these corresponding approaches. We then highlight differences in the assumptions of the three approaches and compare the approaches by using Monte Carlo simulation. We conclude with practical advice for applied researchers regarding the choice of which approach is most closely aligned with their motivating questions, as well as the nature of their data.

## **2. Various Techniques on Longitudinal Reciprocal Influence Investigation**

### **2.1. Autoregressive cross lagged model (ARCL)**

In the existing literature (Bollen & Curran [4]; Rogosa [22]; Rogosa & Willett [23]; Voelkle [26]), ARCL is a common analytic approach to study reciprocal effects between two co-developing constructs. Figure 1 is an illustration of the model with two outcomes with four waves of measurements. In Equations (1),  $x_{it}$  and  $y_{it}$  are modelled by both  $y_{i(t-1)}$  and  $x_{i(t-1)}$ , respectively, where  $i$  indicates the  $i$ -th individual in the sample and  $t$  indicates the  $t$ -th time point ( $t = 2, 3, 4$ ). All random errors,  $\varepsilon_{x_{it}}$ ,  $\varepsilon_{y_{it}}$  are assumed to be independent of each other after accounting for model predictors. Researchers can also assume that autoregressive effects of  $X$  and  $Y$ , and cross lagged effects between the two are constant or time-specific across time points. In the case that autoregressive and cross-lagged effects are assumed to be constant over time, the model is constrained by conditions indicated by Equations (2).



**Figure 1.** ARCL model.

$$x_{it} = \mu_{x_t} + \rho_{1t}x_{i(t-1)} + \rho_{2t}y_{i(t-1)} + \epsilon_{x_{it}}, \quad (1)$$

$$y_{it} = \mu_{y_t} + \rho_{4t}y_{i(t-1)} + \rho_{3t}x_{i(t-1)} + \epsilon_{y_{it}}.$$

$$\rho_{11} = \rho_{12} = \rho_{13} = \rho_1, \quad (2)$$

$$\rho_{21} = \rho_{22} = \rho_{23} = \rho_2,$$

$$\rho_{31} = \rho_{32} = \rho_{33} = \rho_3,$$

$$\rho_{41} = \rho_{42} = \rho_{43} = \rho_4.$$

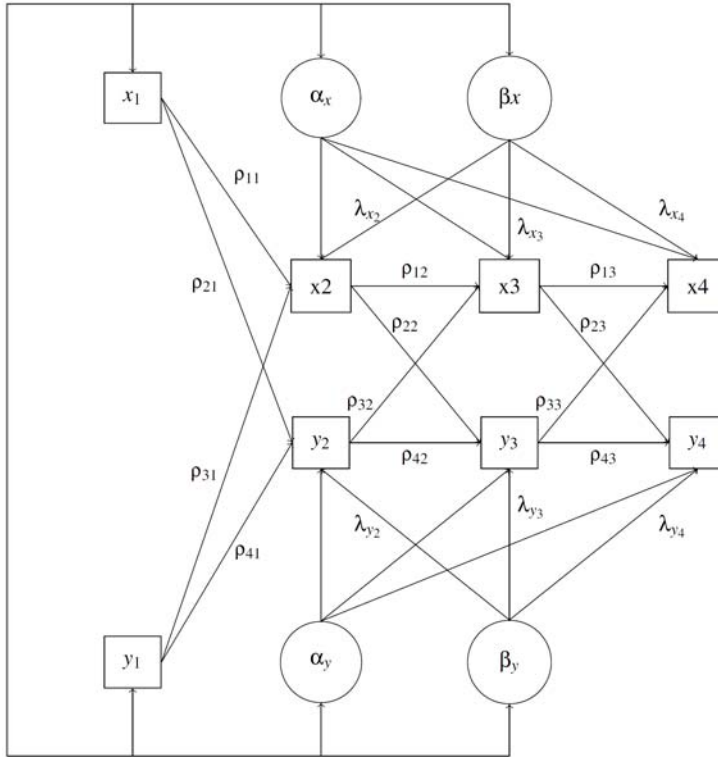
Although ARCL has been used in various studies that investigated co-developing outcomes, its theoretical shortcomings are obvious (Rogosa [22]). By model assumption, residuals are independent of each other, and not correlated with covariates. Analytic results are valid if the strong assumption is satisfied. However, due to the nature of longitudinal data, model residuals of an outcome across different time points may still be correlated due to unique subject effects if no other measures are taken to account for them. Parameter estimates and resulted conclusions may be negatively influenced by the fact that between-and within-subject

variations in model residuals are not separated. Hence, although not widely appreciated by many researchers, the ARCL model, which conflates between-and within-subject variation, does not optimally correspond to within-subject level inferences.

## 2.2. Autoregressive latent trajectory (ALT) model

Bollen and Curran ([4, 8]) integrated growth curve model (GCM) and ARCL model and arrived at an alternative modelling approach, autoregressive latent trajectory (ALT) model. Their approach allows researchers to investigate cross-lagged effects of two longitudinal outcomes, and overall associations of the two outcomes' growth trajectories, while differentiating within-subject random errors from between-subject effects in the form of random effects. Equations (3) illustrate a common ALT model that assumes random intercepts and random slopes  $\xi_{\alpha_{x_i}}$  and  $\xi_{\beta_{x_i}}$ ,  $\xi_{\alpha_{y_i}}$  and  $\xi_{\beta_{y_i}}$  for outcomes  $X, Y$ , respectively.

All random intercepts and slopes are assumed to follow normal distributions independent of model covariates, and the growth rate is modelled by a specific function of time (in this illustration, it is linear with time). Based on individual studies' theoretical frameworks, and number of measurements available in data, researchers can construct their own functional form of latent variables. For example, they may think a model with only random intercepts or one with both linear and quadratic random time slopes fits their theories and data (Bollen & Zimmer [5]). An ALT model imposes a specific functional form of growing trajectory for repeated measures over time. It is very important that the functional form of random effect(s) does not deviate too much from true underlying mechanism so researchers can obtain valid parameter estimates (Voelkle [26]).



**Figure 2.** Random intercept and slope ALT model.

$$x_{it} = \alpha_{x_i} + \lambda_t \beta_{x_i} + \rho_{1t} x_{i(t-1)} + \rho_{2t} y_{i(t-1)} + \varepsilon_{x_{it}}, \quad (3)$$

$$y_{it} = \alpha_{y_i} + \lambda_t \beta_{y_i} + \rho_{4t} y_{i(t-1)} + \rho_{3t} x_{i(t-1)} + \varepsilon_{y_{it}},$$

$$\alpha_{x_i} = \mu_{\alpha_x} + \xi_{\alpha_{x_i}},$$

$$\beta_{x_i} = \mu_{\beta_x} + \xi_{\beta_{x_i}},$$

$$\alpha_{y_i} = \mu_{\alpha_y} + \xi_{\alpha_{y_i}},$$

$$\beta_{y_i} = \mu_{\beta_y} + \xi_{\beta_{y_i}}.$$

### 2.3. Fixed influence cross-lagged (FICL) model

An ALT model has obvious advantages, but still requires a number of modelling assumptions. It is essentially a hybrid approach that carries the random effect model nature of GCM. There are other research methods to separate within-and between-person (unit) variation. From econometricians' perspective, a fixed effect (FE) model is commonly used to analyze longitudinal data that are correlated due to between-subject variations. It sometimes can replace the position of a random effect (RE) model for some unique advantages (Allison [2]). For a detailed discussion of the method, please see Wooldridge [27]. A FE model is a technique to analyze clustered data, including longitudinal data. Compared with the RE model approach, an obvious distinction of FE is that it does not make assumptions on how to model between-subject variations and imposes no assumption that between-subject effects are independent of model predictors. Instead, it assumes a time-invariant effect of each subject's constant characteristics on outcome measurements and conditions this effect out in the process of estimating other time-variant model predictors. Also, in non-experimental studies, researchers often worry about possible model misspecifications due to unobserved factors. FE has the capability of controlling time-invariant influences from all time-invariant unobserved variables.

Assumptions about between-subject variations when applying an ALT model is essential in assuring validity of analytic results, but an approach that combines an ARCL and FE model may avoid making these assumptions. This alternative analytic approach, the fixed influence cross lagged (FICL) model, keeps the study focus on autoregressive and cross lagged parameters like an ARCL and ALT model do and also account for between-subject variations without making ALT's key modelling assumptions of random effects' correct functional form and their independence of other model covariates. Figure 3 is a graphic illustration of a FICL model with four waves of data. In Equations (4), a fixed effect



$\gamma_{x_k}$  ( $k = 2, 3, \dots, n$ ) is imposed on each individual and  $I_k$  is a dichotomous variable that indicates whether a data observation is from the  $k$ -th individual. All autoregressive and cross-lagged effects remain the same as in an ARCL and ALT model.

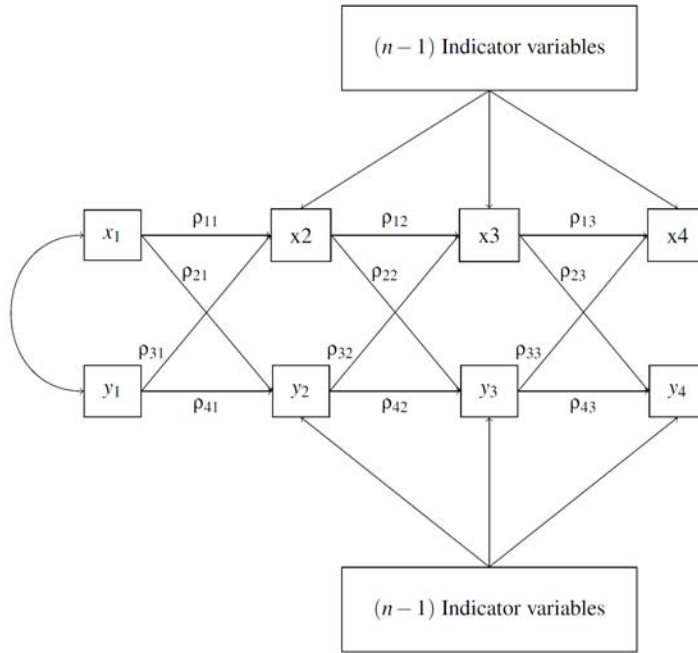


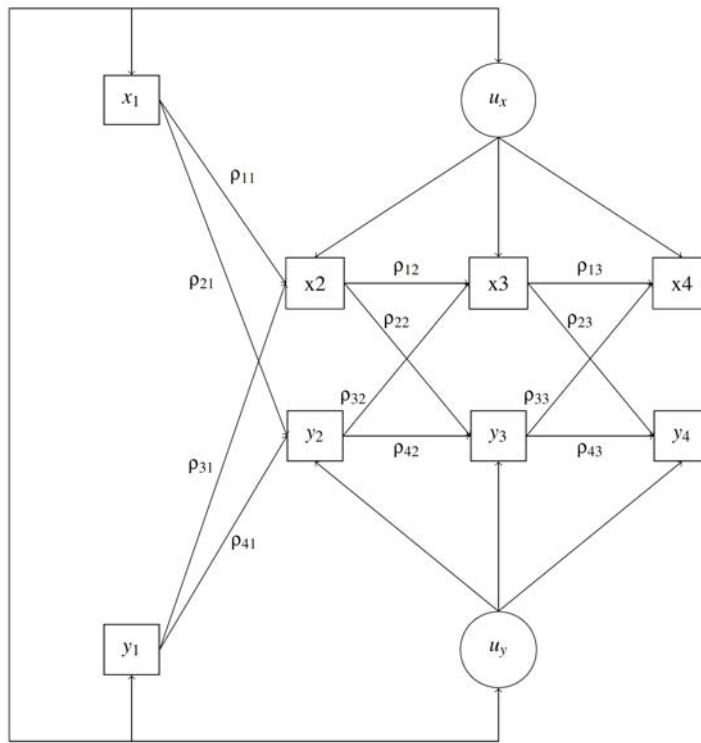
Figure 3. FICL model.

$$x_{it} = \mu_x + \sum_{k=2}^n \gamma_{x_k} I_k + \rho_{1t} x_{i(t-1)} + \rho_{3t} y_{i(t-1)} + \varepsilon_{x_{it}}, \tag{4}$$

$$y_{it} = \mu_y + \sum_{k=2}^n \gamma_{y_k} I_k + \rho_{4t} y_{i(t-1)} + \rho_{2t} x_{i(t-1)} + \varepsilon_{y_{it}}.$$

The comparison of ALT and FICL can be analogized as comparing a RE and FE model. We use an intercept-only ALT model to explain the point (Equations (5)). In Figure 4, subject-specific effects are assumed to be sampled from a larger population that follows normal distribution.

Another important assumption is that  $\xi_{\alpha_{x_i}}$  and  $\xi_{\alpha_{y_i}}$  are independent of  $x_{i(t-1)}$  and  $y_{i(t-1)}$  ( $t = 2, 3, 4$ ). We think the advantages of a FE model over RE model should also apply in the comparison between a FICL and ALT model. First, FICL accounts for between-subject variations without making explicit modelling assumptions about subject specific effects. Second, FICL does not make the assumption that between-subject effects are independent of model covariates like ALT does. Hence, in situations where the assumptions above are not supported, the FICL approach may help researchers avoid corresponding model misspecifications and result in more valid results.



**Figure 4.** Random intercept only ALT model.

$$x_{it} = \alpha_{x_i} + \lambda_t \beta_x + \rho_{1t} x_{i(t-1)} + \rho_{2t} y_{i(t-1)} + \varepsilon_{x_{it}}, \quad (5)$$

$$y_{it} = \alpha_{y_i} + \lambda_t \beta_y + \rho_{4t} y_{i(t-1)} + \rho_{3t} x_{i(t-1)} + \varepsilon_{y_{it}},$$

$$\alpha_{x_i} = \mu_{\alpha_x} + \xi_{\alpha_{x_i}},$$

$$\alpha_{y_i} = \mu_{\alpha_y} + \xi_{\alpha_{y_i}}.$$

### 3. Comparison Among Three Models in an Empirical Example

In order to compare the utilities of the ARCL, ALT, and FICL approach in estimating autoregressive and cross-lagged effects in two comorbid developmental phenomena, we first apply the three models to an empirical dataset collected in Durham Child and Health Development Study (DCHD). The study was initiated in 2002 and 206 young families with healthy, full-term infants were recruited at 3 months of age. The children's ADHD and conduct disorders (CD) symptoms were measured at 18, 24, 30, and 36 months using the ADHD and Oppositional Defiant Disorder scales in Achenbach System of Empirically Based Assessment, Preschool Forms (Achenbach & Rescorla [1]). A subset sample of 159 children who had complete data on the eight data points of interest were included in this illustration. Psychopathology researchers have found that symptoms of the two disorders tend to co-exist (Holmes et al. [14]). However, the underlying mechanisms of how they influence each other are still not well understood (Thapar et al. [25]). We are interested if either condition is the driving force for the development of the other across time. Thus, the three models can provide direct parameter estimations to answer this research question. To simplify this demonstration, we do not go through a potentially complicated model selection process to pick an ALT model with the most appropriate functional form for between-subject random effects, but choose to use a simple intercept only model. Also, we constrain the four sets of autoregressive and cross-lagged effects to be constant over time. All three models are analyzed by MPlus Version 5.0 (Muthén & Muthén [20]).

The key analytic results of autoregressive and cross lagged parameter estimates and their standard errors are presented in Table 1. As we can see, while three models result in comparable autoregressive effect estimates ( $\hat{\rho}_1, \hat{\rho}_4$ ), the ARCL model produces cross-lagged effect estimates ( $\hat{\rho}_2, \hat{\rho}_3$ ) that are in the opposite direction of those from the ALT and FICL model. Also, although the direction of the cross-lagged effect estimates by the ALT and FICL method are consistent with each other, there are considerable differences in the magnitude of the effects. Due to the fact that the three models produce very different estimates on parameters of interest, choosing an analytic model that is consistent with a theoretical framework and appropriate for specific data situation can have a big impact on the quality of potential study results. Since the interest of the example analysis is to investigate within-person autoregressive and cross lagged associations between ADHD and CD symptoms over time and the ARCL model does not differentiate the within-person random errors and between-person effects, ALT and FICL model approach should be chosen over ARCL model under this situation.

**Table 1.** Comparison of the models in an empirical example

Parameter estimates	ARCL	ALT	FICL
$\hat{\rho}_1$	.441(.043)	.471(.064)	.442(.095)
$\hat{\rho}_2$	.199(.048)	-.127(.066)	-.321(.08)
$\hat{\rho}_3$	.081(.039)	-.13(.054)	-.318(.071)
$\hat{\rho}_4$	.483(.043)	.377(.087)	.382(.104)

It is more difficult to decide preference between ALT and FICL based on these empirical results because it is hard to know if modelling assumptions of an ALT or FICL model is violated. Based on our experience with RE and FE models, if an intercept-only RE model is valid, changing to FE model approach would produce similar model fixed effect estimates in general. We expect that this observation can be applied here as well. As a result, the discrepancies in cross-lagged effect estimates between the ALT and FICL approach cause us to have some reservations on our confidence in the ALT model under this situation.

The selection between a RE model and FE model can also be facilitated by the Hausman test in practice (Hausman [13]). The between-subject effects are assumed to be normally distributed random variables independent of other model covariates in RE models, whereas time-invariant between-subject effects are considered to be fixed in FE models. When the assumptions of RE models are satisfied, parameter estimates are more efficient than those from corresponding FE models. However, when the assumptions are unlikely to hold, estimates from FE models are less biased than the ones from RE models. Let  $\hat{\beta}_{RE}$  and  $\hat{\beta}_{FE}$  be the coefficient estimates of a RE model and FE model, respectively. The variance of the difference between  $\hat{\beta}_{RE}$  and  $\hat{\beta}_{FE}$  is denoted as  $\sum_{\hat{\beta}_{RE}-\hat{\beta}_{FE}}$ . Under the null hypothesis that the individual effects are identically and independently distributed in the RE model specification, the followings hold:

$$(\hat{\beta}_{RE} - \hat{\beta}_{FE})' \sum_{\hat{\beta}_{RE}-\hat{\beta}_{FE}}^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE}) \sim \chi_k^2 \text{ (Greene [11]; Hausman [13]).}$$

When the null hypothesis is rejected by the asymptotic  $\chi^2$  test, the FE model should be favored. Since the ALT and FICL model can be thought of as an efficient estimation approach to study autoregressive and cross-lagged associations between two longitudinal outcomes and another alternative approach without the assumption that time-invariant between-subject effects are uncorrelated with other model covariates, we can use Hausman's test for model selection in this situation. In our example, we conduct the test on the autoregressive and cross-lagged effect estimates  $(\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3, \hat{\rho}_4)$ . The resulted Hausman test ( $\chi^2$ ) statistic with 4 degrees of freedom is 8.932,  $p < .05$ . The test result indicates that the FICL model should be chosen over the intercept-only ALT model in this empirical example.

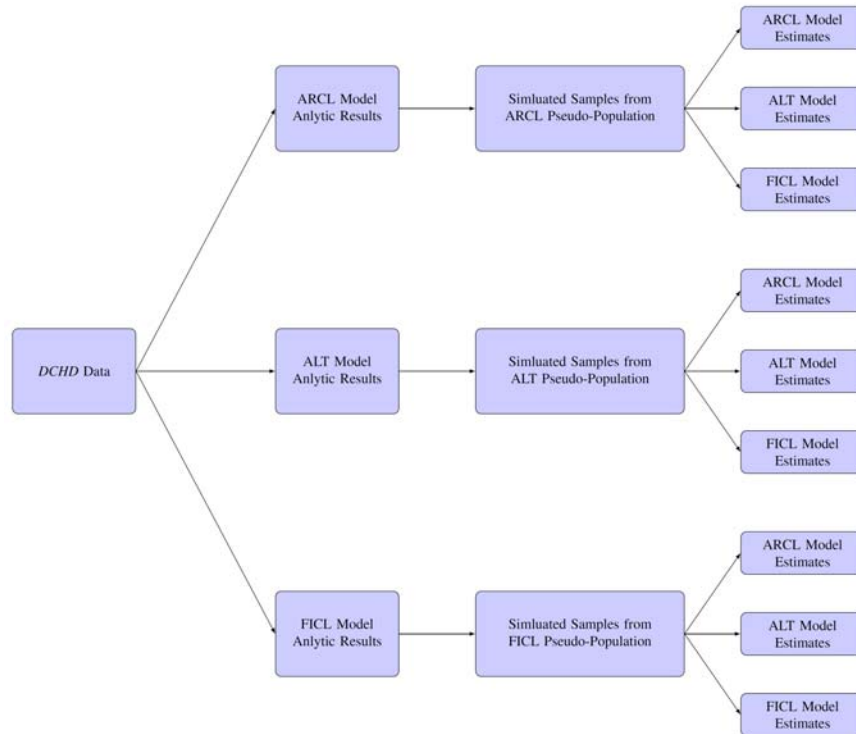
## 4. Monte Carlo Simulation

### 4.1. Data-driven resampling framework

In order to further compare the three approaches and explore possible advantages of FICL over the other two approaches, we decide to conduct a Monte Carlo simulation study by following the method of “sampling study” (MacCallum et al. [18], p. 495). The sampling study is an observed-data-driven repeated sampling approach for stability check and cross-validation of model specification. In MacCallum and colleagues’ study, they randomly drew series of data subsets with varying sample sizes from the observed data reservoir, which was treated as the finite population. Under each sample size, many pairs of subset of data are used to obtain the cross-validation indices, model fit indices and modification indices. Comparing the quantities obtained from the re-sampled data with the “true values” obtained from the finite population, the researchers can evaluate how the specific-model’s stability has been affected by the simulation factor (i.e., sample size). We adapt the resampling framework in our simulation study. Instead of varying sample sizes in various samples drawn, we use different sets of parameter values to represent various possible underlying true models and compare performance of the ARCL, ALT, and FICL approach across various situations.

### 4.2. Simulation design

The simulation design of the current study is closely based on the analysis results of the empirical example. Figure 5 depicts the protocol of our study including the following simulation procedures.



**Figure 5.** Simulation study procedure.

**Step 1.** Since researchers do not know what a true model for their data is, we first use each of the three models and corresponding model-based estimates from analyzing the actual DCHD data as a “true” model.

**Step 2.** We simulate data using each of the three “true” models. For example, we generate data based the ARCL model using the relevant parameter values as the pseudo-population parameter values, and the resulted data are considered to be from a hypothetical true relationship between two co-developing phenomena.

**Step 3.** In this step, we analyze each “true” model’s data with the ARCL, ALT, and FICL model and compared the analytic results with true parameter values.

**Step 4.** After obtaining all the simulation results, cross-validation analyses verify the model selection in the final step.

### 4.3. Simulation study hypotheses

Based on our discussion, we have the following three hypotheses: (1) the ARCL method is unlikely to have any obvious advantage over the other two approaches; (2) ALT and FICL methods are comparable to each other if the model assumptions for ALT are likely to hold, and both are better than ARCL; and (3) a FICL model should be advantageous over an ALT model if ALT model assumptions are not satisfied. As a result, we expect to get similar analytic results across three methods if the true underlying process is consistent with ARCL. Also, results from the ALT and FICL model are superior to those from ARCL approach if the underlying process is ALT, and the former two are comparable to each other. Finally, analysis using FICL model should outperform the other two methods if the true process is FICL.

### 4.4. Simulation replications

In our simulation, 2,000 samples of size 200 observations are randomly drawn based on each hypothetical true model. Each observation contains eight variables of interest, that is, four repeated measures on each of two codeveloping phenomena. ARCL, ALT, and FICL models are fit on sample  $r$  to obtain parameter estimates ( $r = 1, 2, \dots, 2,000$ ). The means and sample standard deviations of parameter estimates are recorded for later comparison purpose. Since the motivation of our study is to help researchers probe autoregressive and cross lagged effects of two reciprocal phenomena, evaluations of different analysis approaches on simulated data are based on bias and mean-squared error (MSE) of the estimates of interest. Let  $\hat{\rho}_i$  be the estimator of the  $i$ -th pseudo-population parameter  $\rho_i$  ( $i = 1, 2, 3, 4$ ). We compute  $\text{Bias}(\hat{\rho}_i) = (\bar{\hat{\rho}}_i - \rho_i)$ , and  $\text{MSE}(\hat{\rho}_i) = [\text{Bias}(\hat{\rho}_i)]^2 + \hat{\text{Var}}(\hat{\rho}_i)$ .



#### 4.5. Simulation results

*ARCL as the True Model.* Table 2 presents analytic results on 2,000 simulated data based on ARCL model. With no surprise, we find that if the data is simulated based on the ARCL model, using the model itself to estimate parameters of interest is the best choice resulting the smallest empirical biases in the coefficient estimates and MSE. However, the performance of the other two approaches is largely comparable to the hypothetical true model. Only  $\rho_3$  may be concluded to not differ from 0 significantly; all other statistical inference conclusions are consistent across the three approaches. This set of simulation study results is consistent with our first hypothesis that ARCL model is unlikely to have significant advantages over ALT or FICL even if ARCL represents the underlying reciprocal process.

**Table 2.** Comparison of the models using Monte Carlo simulation with ARCL

Analytic model	Parameter estimator	True parameter value	Mean	S.D.	Bias	MSE
ARCL	$\hat{\rho}_1$	0.441	0.437	0.039	-0.004	0.0015
	$\hat{\rho}_2$	0.199	0.199	0.044	0	0.0019
	$\hat{\rho}_3$	0.081	0.081	0.035	0	0.0012
	$\hat{\rho}_4$	0.483	0.481	0.039	-0.002	0.0015
ALT	$\hat{\rho}_1$	0.441	0.439	0.062	-0.002	0.0038
	$\hat{\rho}_2$	0.199	0.193	0.053	-0.006	0.0028
	$\hat{\rho}_3$	0.081	0.063	0.041	-0.018	0.0020
	$\hat{\rho}_4$	0.483	0.504	0.061	0.021	0.0042
FICL	$\hat{\rho}_1$	0.441	0.436	0.104	-0.005	0.0108
	$\hat{\rho}_2$	0.199	0.190	0.078	-0.009	0.0062
	$\hat{\rho}_3$	0.081	0.052	0.062	-0.029	0.0047
	$\hat{\rho}_4$	0.483	0.533	0.121	0.050	0.0171

*ALT as the True Model.* Table 3 contains analytic results of 2,000 simulations based on the ALT model. As we can see, when the underlying process conforms with ALT, analytic results using the ALT and FICL approach are very close in general with FICL approach resulting more conservative standard errors. All final statistical inference using the two methods are exactly the same. However, ARCL model results in totally different estimates for cross lagged effects ( $\rho_2$  and  $\rho_3$ ) with the wrong direction. As a result, conclusions of the cross lagged effects would be incorrect if researchers use the ARCL method when the underlying true model is ALT under this data scenario. Hence, these results are aligned with our second hypothesis that ALT and FICL methods are comparable to each other if the model assumptions for ALT seem to hold, and both are better than ARCL.

**Table 3.** Comparison of the models using Monte Carlo simulation with ALT

Analytic model	Parameter estimator	True parameter value	Mean	S.D.	Bias	MSE
ARCL	$\hat{\rho}_1$	0.471	0.421	0.039	-0.050	0.0040
	$\hat{\rho}_2$	-0.127	0.144	0.044	0.271	0.0754
	$\hat{\rho}_3$	-0.130	0.036	0.039	0.166	0.0291
	$\hat{\rho}_4$	0.377	0.447	0.043	0.070	0.0067
ALT	$\hat{\rho}_1$	0.471	0.470	0.046	-0.001	0.0021
	$\hat{\rho}_2$	-0.127	-0.124	0.046	0.003	0.0021
	$\hat{\rho}_3$	-0.130	-0.130	0.041	0	0.0017
	$\hat{\rho}_4$	0.377	0.377	0.058	0	0.0034
FICL	$\hat{\rho}_1$	0.471	0.473	0.087	0.002	0.0076
	$\hat{\rho}_2$	-0.127	-0.126	0.069	0.001	0.0048
	$\hat{\rho}_3$	-0.130	-0.131	0.057	-0.001	0.0033
	$\hat{\rho}_4$	0.377	0.379	0.082	0.002	0.0067

*FICL as the True Model.* Finally, we present the results of analyzing 2,000 samples drawn from FICL model in Table 4. Under this setup, we see clear advantage of using the FICL model to analyze data over the other two approaches. Again, the ARCL model results in totally wrong cross-lagged effect estimates. Although the ALT model estimates the two cross-lagged effects to be in the consistent direction of the true parameter values, the magnitude of the regression coefficients differs by roughly 60%. We conclude that the final set of results confirms our third hypothesis that FICL clearly outperforms the other two methods if the true reciprocal process between two outcomes of interest conforms with the model.

**Table 4.** Comparison of the models using Monte Carlo simulation with FICL

Analytic model	Parameter estimator	True parameter value	Mean	S.D.	Bias	MSE
ARCL	$\hat{\rho}_1$	0.442	0.472	0.039	0.030	0.0024
	$\hat{\rho}_2$	-0.321	0.168	0.041	0.489	0.2408
	$\hat{\rho}_3$	-0.318	0.072	0.036	0.390	0.1534
	$\hat{\rho}_4$	0.382	0.497	0.043	0.115	0.0151
ALT	$\hat{\rho}_1$	0.442	0.487	0.084	0.045	0.0091
	$\hat{\rho}_2$	-0.321	-0.103	0.097	0.218	0.0569
	$\hat{\rho}_3$	-0.318	-0.119	0.070	0.199	0.0445
	$\hat{\rho}_4$	0.382	0.396	0.135	0.014	0.0184
FICL	$\hat{\rho}_1$	0.442	0.445	0.097	0.003	0.0094
	$\hat{\rho}_2$	-0.321	-0.321	0.074	0	0.0055
	$\hat{\rho}_3$	-0.318	-0.319	0.062	-0.001	0.0038
	$\hat{\rho}_4$	0.382	0.388	0.107	0.006	0.0115

## 5. Discussion

In this article, we have briefly reviewed and compared three common analytic techniques that can be used to inform time-linked associations between two constructs measured at a number of time points: ARCL, ALT, and FICL model. Under the restricted assumption that the random errors in repeated measures are conditionally independent of each other after controlling for earlier measurements and other model covariates, ARCL model provides valid estimates of autoregressive cross lagged associations. However, ALT and FICL approach seem to have close performance to ARCL method and lead to same valid conclusions, as shown in our simulation study. ALT and FICL method relax that strict ARCL assumption by acknowledging that outcome residuals may still be correlated after controlling for earlier measurements in the form of random effects and fixed effects. As a result, in many situations, ALT and FICL approach may be more appealing to researchers who are not comfortable with the model residual assumption imposed by an ARCL model.

The main difference between ALT and FICL model is their ways to parameterize between-subject effects. In an ALT model, between-subject effects are assumed to be random and have clearly defined functional forms. They are considered to be normally distributed and independent of other model covariates. In FICL models, time-invariant between-subject effects are considered to be fixed for each subject and no assumptions of distribution or independence of model covariates are imposed. When assumptions for an ALT model is satisfied, using ALT and FICL may arrive at close results, but the former can have smaller standard errors for interested parameter estimates than the latter. The loss of efficiency by FICL does not have an impact on study conclusions in our specific simulations, but it is possible in some cases that statistical significant conclusions are resulted by ALT and insignificant ones by FICL. Also, an ALT model accommodates time-variant between-subject effects better

than a FICL model by allowing estimation of random effects of time-variant variables. In this paper, we only compare the performance between FICL and random-intercept-only ALT model. Comparisons between performance by a FICL model and ALT models with more complicated functional form of random effects may be conducted in future studies.

When ALT model assumptions are violated, we have shown with simulations that a FICL model results in less biased parameter estimates than an ALT model, and has the best performance out of the three approaches. Like a FE model, a FICL model incorporates all time-invariant fixed effect into one single subject-specific effect. This property allows the FICL approach to be able to control for constant effects from important time-invariant observed variables, although it also means the model is incapable of estimating effects of time-invariant covariates on outcome variables like an ARCL or an ALT model does. The Hausman test should be conducted to facilitate the selection between an ALT and FICL model.

In summary, all three approaches that investigate autoregressive cross-lagged effects have appropriate applications. It is beneficial for researchers to be aware of the key model assumptions behind these methods, understand in what situations each model may produce valid results, and pick an appropriate approach based on their specific theoretical framework, understanding of their data, and formal model selection tests (e.g., Hausman test).

### References

- [1] T. Achenbach and L. Rescorla, *Manual for Aseba Preschool Forms and Profiles*, Burlington, VT: Research Center for Children, Youth, and Families, University of Vermont, 2000.
- [2] P. Allison, *Fixed Effects Regression Methods for Longitudinal Data using SAS*, Cary, NC: SAS Publishing, 2005.
- [3] P. Allison, *Fixed Effects Regression Models*, Sage Publications, London, 2009.

- [4] K. Bollen and P. Curran, Autoregressive latent trajectory (ALT) models a synthesis of two traditions, *Sociological Methods & Research* 32(3) (2004), 336-383.
- [5] K. Bollen and C. Zimmer, An Overview of the Autoregressive Latent Trajectory (ALT) Model, In J. van Montfort, J. Oud & A. Satorra (Eds.), *Longitudinal Research with Latent Variables*, Springer, New York, NY, 2010.
- [6] A. Chan and F. Poulin, Monthly instability in early adolescent friendship networks and depressive symptoms, *Social Development* 18(1) (2009), 1-23.
- [7] J. Cullum, S. Armeli and H. Tennen, Drinking norm-behavior association over time using retrospective and daily measures, *Journal of Studies on Alcohol and Drugs* 71(5) (2010), 769-777.
- [8] P. Curran and K. Bollen, The best of both worlds: Combining autoregressive and latent curve models, *New Methods for the Analysis of Change* (2001), 107-135.
- [9] A. Farrell, Structural equation modeling with longitudinal data: Strategies for examining group differences and reciprocal relationships, *Journal of Consulting and Clinical Psychology* 62(3) (1994), 477-487.
- [10] E. Ferrer and J. McArdle, Longitudinal modeling of developmental changes in psychological research, *Current Directions in Psychological Science* 19(3) (2010), 149-154.
- [11] W. Greene, *Econometric Analysis* (5th Ed.), Upper Saddle River, Prentice Hall, NJ, 2003.
- [12] K. Grimm, Multivariate longitudinal methods for studying developmental relationships between depression and academic achievement, *International Journal of Behavioral Development* 31(4) (2007), 328-339.
- [13] J. A. Hausman, Specification tests in econometrics, *Econometrica: Journal of the Econometric Society* (1978), 1251-1271.
- [14] J. Holmes, A. Payton, J. Barrett, R. Harrington, P. McGuffin and M. Owen et al., Association of DRD4 in children with ADHD and comorbid conduct problems, *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 114(2) (2002), 150-153.
- [15] N. Kteily, J. Sidanius and S. Levin, Social dominance orientation: Cause or 'mere effect'? Evidence for SDO as a causal predictor of prejudice and discrimination against ethnic and racial out groups, *Journal of Experimental Social Psychology* 47(1) (2010), 208-214.
- [16] K. Luyckx, I. Seiffge-Krenke and S. Hampson, Glycemic control, coping, and internalizing and externalizing symptoms in adolescents with type 1 diabetes: A cross-lagged longitudinal approach a cross-lagged longitudinal approach, *Diabetes Care* 33(7) (2010), 1424-1429.
- [17] R. MacCallum, C. Kim, W. Malarkey and J. Kiecolt-Glaser, Studying multivariate change using multilevel models and latent curve models, *Multivariate Behavioral Research* 32(3) (1997), 215-253.

- [18] R. C. MacCallum, M. Roznowski and L. B. Necowitz, Model modifications in covariance structure analysis: The problem of capitalization on chance, *Psychological Bulletin* 111(3) (1992), 490.
- [19] L. Mayer and S. Carroll, Testing for lagged, cotemporal, and total dependence in cross-lagged panel analysis, *Sociological Methods & Research* 16(2) (1987), 187-217.
- [20] L. Muthén and B. Muthén, *Mplus Version 5.0*, Los Angeles, CA: Muthén & Muthén, 1998-2008.
- [21] H. Phan, Predicting change in epistemological beliefs, reflective thinking and learning styles: A longitudinal study, *British Journal of Educational Psychology* 78(1) (2008), 75-93.
- [22] D. Rogosa, A critique of cross-lagged correlation, *Psychological Bulletin* 88(2) (1980), 245-258.
- [23] D. Rogosa and B. Willett, Satisfying a simplex structure is simpler than it should be, *Journal of Educational Statistics* 10(2) (1985), 99-108.
- [24] A. Sayer and J. Willett, A cross-domain model for growth in adolescent alcohol expectancies, *Multivariate Behavioral Research* 33(4) (1998), 509-543.
- [25] A. Thapar, R. Harrington and P. McGuffin, Examining the comorbidity of ADHD-related behaviours and conduct problems using a twin study design, *The British Journal of Psychiatry* 179(3) (2001), 224-229.
- [26] M. Voelkle, Reconsidering the use of autoregressive latent trajectory (ALT) models, *Multivariate Behavioral Research* 43(4) (2008), 564-591.
- [27] J. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, MA, 2002.

