# One Step Toward Causality: Unobserved Time-Invariant Confounding in Cross-Lagged Panel Models

Pepijn Vink (6100252)[1]

**ARTICLE HISTORY**
Compiled December 12, 2023

[1] Methodology and Statistics for the Behavioral, Biomedical, and Social Sciences, Utrecht University,

**ABSTRACT**
Cross-lagged panel models are a popular analysis technique to analyze reciprocal effects between multiple variables over time. Although these effects are often interpreted as causal, confounders should be adequately accounted for before causality can be inferred. The current paper explores the effect of unobserved time-invariant confounders on estimates of the Random Intercept Cross Lagged Panel Model (RI-CLPM) and the Dynamic Panel Model (DPM), two popular panel models that have been claimed to control for time-invariant confounders. A simulation study shows that when true effects are stable over time, the RI-CLPM and the DPM yield unbiased estimates. When the effects are time-varying, the estimates from both models are biased.

**KEYWORDS**
RI-CLPM; DPM; causality; longitudinal; simulation; confounding

## 1. Introduction

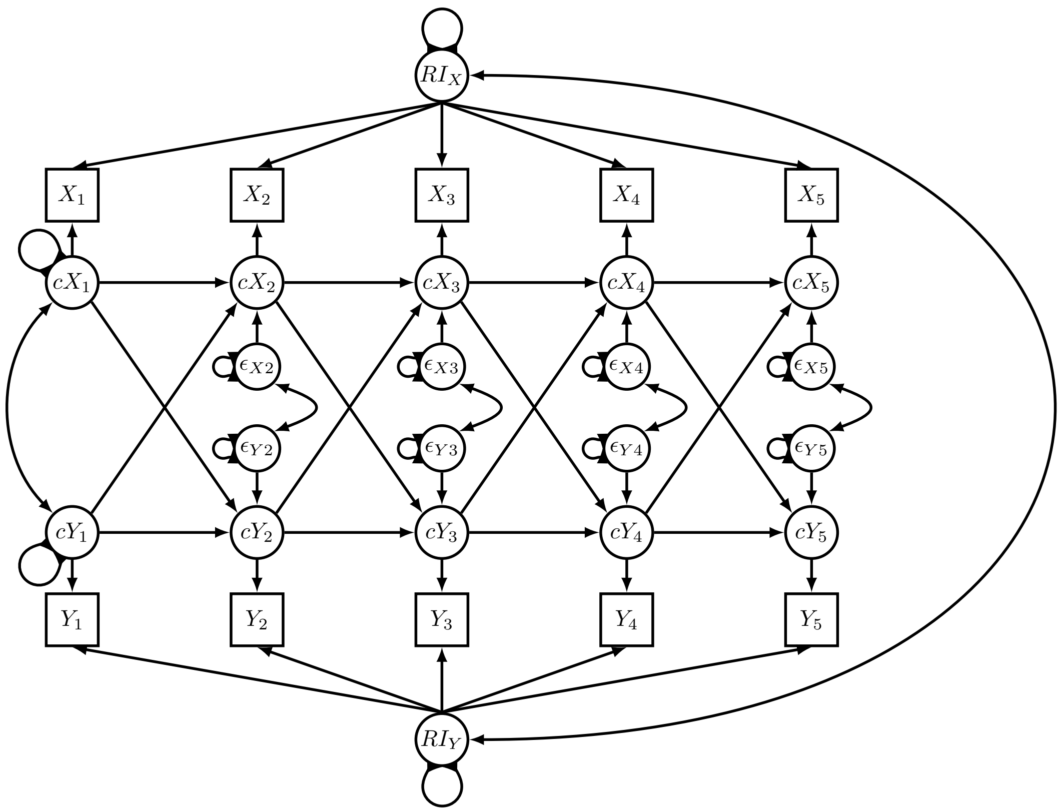
Cross-lagged panel designs are a popular method in psychological research to investigate relationships between two or more variables over time. In the last few years, the random-intercept cross-lagged panel model (RI-CLPM, Hamaker et al., 2015; Mulder & Hamaker, 2021) has become a popular method of analyzing cross-lagged relationships in panel data (longitudinal, non experimental data with 3-10 measurement moments). The RI-CLPM, shown in Figure 1a, is an extension of the cross-lagged panel model (CLPM)

CONTACT: Pepijn Vink (6100252). Email: p.a.vink@uu.nl.

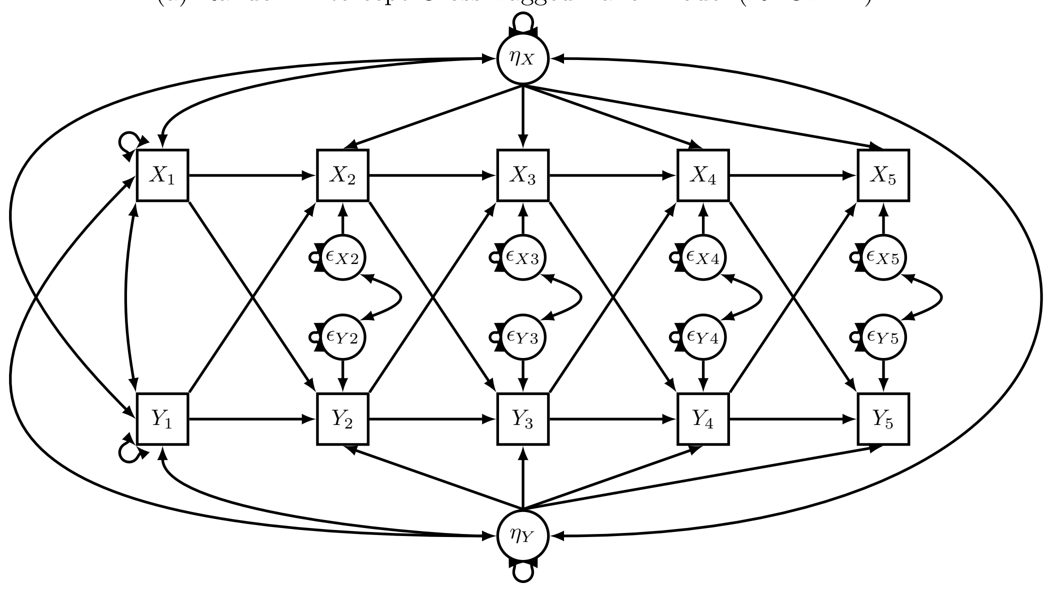
and decomposes observed variables (the boxes) into stable between person differences (the 'between' part of the model) and within person dynamics (the 'within' part of the model). The within person dynamics are usually most of interest. Specifically, the interest is often in the effects of the variables on each other at later timepoints: the cross-lagged effects. These cross-lagged parameters are often (albeit implicitly) interpreted as causal effects. However, one assumption that such causal interpretations rely on is that both time-varying and time-invariant confounders are adequately controlled for.

Usami et al. (2019) show that when certain assumptions are met, the random intercept in the RI-CLPM controls for unobserved heterogeneity. This requires, in particular, that the effects of the confounders on the variables of interest are stable over time. However, it has not yet been studied how the RI-CLPM performs when the effects of unobserved time-invariant confounders are of a time-varying nature. Murayama & Gfrörer (2022) show that the Dynamic Panel Model (DPM), a related model to the RI-CLPM and shown in Figure 1b, may be an alternative in this case, but the downside is that this model does not separate within from between effects. Furthermore, it is unknown how either of these models perform when the underlying causal model is more complex, for example when multiple confounders with time-varying effects are involved.

Therefore, the effect of time-invariant confounders on both the RI-CLPM and the DPM should be explored further. The current research report will evaluate the extent to which unobserved time-invariant confounders affect the estimates of the RI-CLPM and DPM. I will explore different types of time-invariant confounders in panel data (i.e. with time-stable effects versus time-varying effects) and assess their effects on estimates of the RI-CLPM and DPM when they are unobserved. This will serve as a preliminary exploration for the thesis, which will address the use of causal inference




(a) Random Intercept Cross-Lagged Panel Model (RI-CLPM)




(b) Dynamic Panel Model (DPM)

Figure 1. Two Popular Models in Panel Research. Boxes Indicate Observed Variables and Circles Indicate Latent Variables.

methods such as propensity score adjustment and inverse probability weighting (e.g. Brown et al., 2021; Vansteelandt & Daniel, 2014) to control for observed time-invariant confounders in the RI-CLPM.

This report is structured as follows. I will start with a conceptual comparison between the RI-CLPM and the DPM. Then, a hypothetical data generating mechanism that includes multiple time-invariant confounders will be introduced and the potential of the RI-CLPM and the DPM to control for unobserved confounding when this mechanism is true will be discussed. After this, a simulation study will be performed to assess the performance of the RI-CLPM and the DPM under the data generating mechanism. Results and implications will be discussed.

## 2. A Comparison of the RI-CLPM and the DPM

The RI-CLPM (Hamaker et al., 2015) is often used with the goal to separate within person dynamics from stable between person differences. It decomposes the observed scores of an individual in a stable person mean (the random intercept representing the 'between' part of the model) and a temporal deviation from that mean (the 'within' part of the model). The model implies only direct effects of the stable trait as well as no effect of the stable trait of one variable on the observed scores of the other (other than through its covariance with the other stable trait factor).

The DPM, however, is usually used when the goal is to model lagged effects while using a latent factor to control for time-invariant confounders. It does not explicitly decompose within person dynamics from stable between person differences, and regresses the observed scores on each other. The latent factors in the DPM are sometimes called 'accumulating factors' (Usami et al., 2019), as their effects on the observed variables are both direct, as well as indirect through lagged relationships between the observed variables themselves. Furthermore, to account for the fact that measurements are usually sampled at a random moment in time in an ongoing process, the loading for the first timepoint is estimated freely (Hamaker, 2005), which is not necessary in the RI-CLPM.

Hamaker (2005) shows that in specific cases, the RI-CLPM and the DPM are statistically equivalent and yield equivalent estimates of the lagged parameters. Specifically, this is the case when lagged parameters are stable over time and the covariances in the DPM at the first timepoint are constrained to reflect this (rather than estimated freely). This also implies that when these conditions hold in a causal model, both the RI-CLPM and the DPM should recover the true effects.

## 3. Example of a Data Generating Mechanism

Consider the causal Directed Acyclic Graph (DAG) in figure Figure 2. It shows a dynamic process for $t = 1, ..., 5$ between time-varying variables $X$ and $Y$ and includes time-invariant confounders $C_1$ and $C_2$. This causal DAG does not yet have parametric assumptions. This DAG is most similar to the dynamic panel model, as observed values are determined by observed values at previous timepoints, but it can also be interpreted as only explicitely modeling the within person process.

Assuming the underlying process has been going on for long enough that it has stabilized around an equilibrium, when lagged effects as well as the effects of the confounders are time-stable, in both the RI-CLPM and the DPM the latent factor will be a linear combination of the confounders (Usami et al., 2019). Therefore, when these conditions hold, both models should already yield unbiased estimates of the lagged effects, even when the confounders are unobserved. However, when the effects of the

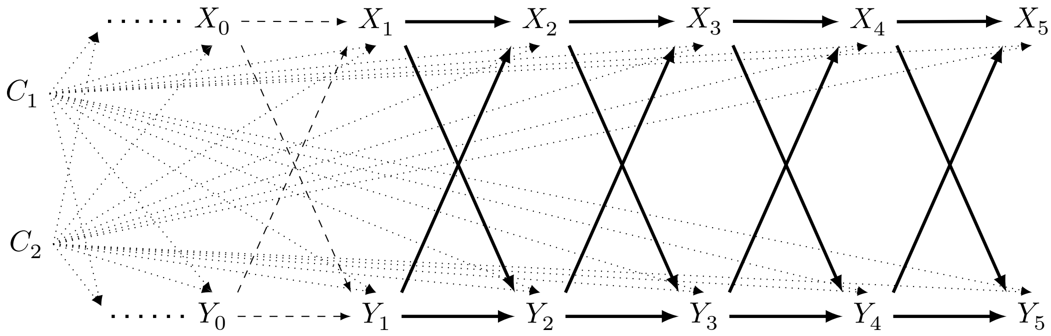



Figure 2. Directed Acyclic Graph for Cross-Lagged Relationships with 2 Confounders. Dashed Lines Indicate the Process Before the Observed Timepoints.

confounders are not time-stable, the models may not (fully) take the confounders into account thus resulting in biased effects. In the next section, a simulation study will be described that simulates from our DAG and evaluates the performance of the RI-CLPM and the DPM when confounders are unobserved.

## 4. Methods

In previous sections, the RI-CLPM and the DPM have been compared conceptually and their ability to account for unobserved time-invariant confounding has been touched upon. In this section, a simulation study will be presented that evaluates their performance under three different scenarios. These scenarios are based on the data generating mechanism in figure Figure 2 and differ with respect to the stability of the effects of the confounders

### 4.1. *Data Generation*

Below, the different scenarios of the data generating mechanism, that will be simulated are described, each having different effects of the confounders. These all follow the following general data generating mechanism:

For each person $i$ at timepoint $t$,

$$x_{it} = \phi_{xx}x_{i,t-1} + \phi_{xy}y_{i,t-1} + \gamma_{1t}C_{1i} + \gamma_{2t}C_{2i} + \epsilon_{xit},$$

$$y_{it} = \phi_{yy}y_{i,t-1} + \phi_{yx}x_{i,t-1} + \delta_{c1t}C_{i} + \delta_{2t}C_{2i} + \epsilon_{yit}.$$

Furthermore, at the first sampled timepoint,

$$x_{i} = \gamma_{1}C_{1i} + \gamma_{2}C_{2i} + \epsilon_{xi},$$

$$y_{i} = \delta_{1}C_{1i} + \delta_{2}C_{2i} + \epsilon_{yi}.$$

In addition,

$$\epsilon_{xt} \sim \mathcal{N}(0, \psi_{x}),$$

$$\epsilon_{yt} \sim \mathcal{N}(0, \psi_{y}),$$

$$C_{1} \sim \mathcal{N}(0, \psi_{C_1}),$$

$$C_2 \sim \mathcal{N}(0, \psi_{C_2}).$$

All lagged effect (cross-lagged effects and autoregressions) and residual variances will be time-stable, whereas effects of the confounders may be time-varying. For all scenarios lagged effects are set to $\phi_{xx} = 0.2$, $\phi_{yy} = 0.3$, $\phi_{xy} = 0.15$, $\phi_{yx} = 0.1$ (based on Mulder, 2023) and the residual variances are set to $\psi_x = 1$, $\psi_y = 1$, $\psi_{C_1} = 1$, and $\psi_{C_2} = 1$.

Three scenarios will be simulated. For all scenarios, at the start, $\gamma_1 = 0.3$, $\gamma_2 = 0.8$, $\delta_1 = 0.5$, and $\delta_2 = 0.2$. For scenario 1, the effects of the confounders remain stable. For scenario 2, at $t = 3$, the effect of $C_1$ on $x$ and $y$ decreases and remains stable afterwards, $\gamma_{1t} = 0.1$ and $\delta_{1t} = 0.2$. For scenario 3, at $t = 3$ all effects confounder effects change and afterwards remain stable $\gamma_{1t} = 0.6$ and $\delta_{1t} = 0.2$, $\gamma_{2t} = 0.3$ and $\delta_{2t} = 0.5$.

1000 datasets are simulated according to each scenario with N=500. To allow for convergence to an equilibrium, 50 timepoints are simulated, of which 45 are used as burn-in. The remaining 5 timepoints are used for analysis.

### 4.2. *Analysis*

To assess the performance of the described models when the confounders are unobserved, all simulated datasets described above will be analyzed using the RI-CLPM and the DPM, as well as versions of these models with free factor loadings, as these may, in part, capture some time-varying heterogeineity due to the confounders. Models that do not converge or do not result in positive definite covariance matrices will be excluded. All simulation will be done using R (R Core Team, 2022) and models will be fit using the lavaan package (Rosseel, 2012).

## 5. Results

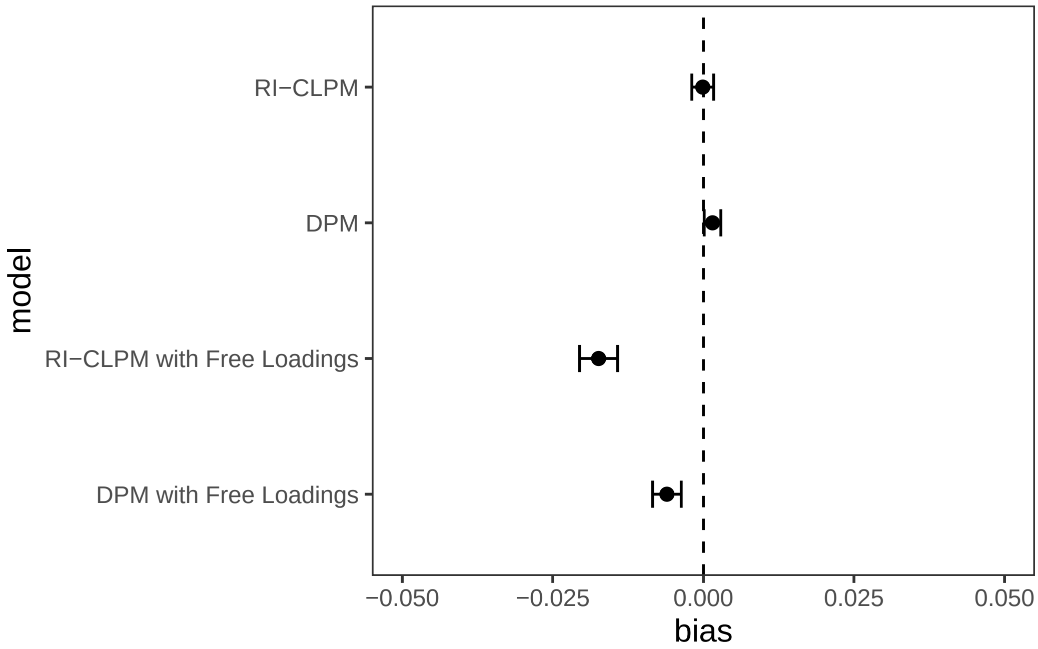
Figure 3 shows the biases for the $x_4 \rightarrow y_5$ effect in scenario 1 and 2. The error bars indicate the Monte Carlo error, which can be interpreted as the standard error for the bias estimates (Morris et al., 2019). Scenario 3 is not included in the results, because all RI-CLPM models and almost all DPM models resulted in a non positive-definite covariance matrix of the latent variables.

Figure 3a shows that under scenario 1, as expected, both the RI-CLPM and the DPM yield unbiased estimates. Surprisingly, freeing the factor loadings, however, results in bias. Specifically, they yield a negative bias, indicating that they underestimate the effect. Furthermore, the RI-CLPM has an MSE of 0.0032663 and the DPM has an MSE of 0.0019128, indicating that under this scenario, the DPM may be a better choice.

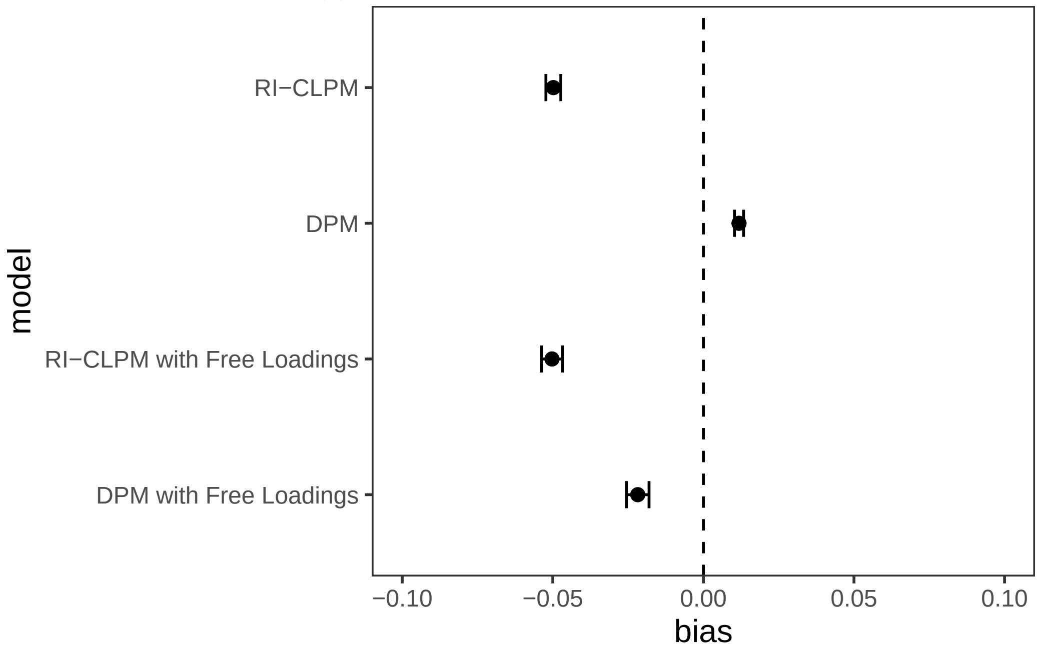
Under scenario 2, where the effects one of the confounders is time-varying, both the RI-CLPM and the DPM yield biased effects. Furthermore, freeing the factor loadings does not seem to capture the time-varying effects of the confounders. Furthermore, the RI-CLPM has an MSE of 0.0060792 and the DPM has an MSE of 0.0019639 under this scenario.

## 6. Discussion

Pearl & Mackenzie (2019) argue that science has been in a 'causal revolution' and as Haber et al. (2022) shows, there is an abundance of causal language present in modern research. It is therefore essential that causal effects are estimated with great care.




(a) Scenario 1: Time-Stable Effects




(b) Scenario 2: Time-Varying Effects for C1, Time-Stable Effects for C2

Figure 3. Biases for Scenario 1 and Scenario 2. The dots indicate the bias of the models and the error bars indicate the Monte Carlo Error

Evaluating causality using reciprocal effects in longitudinal data can bring researchers in the social science one step closer to obtaining the 'true' causal effect, but more will be necessary.

It has been known that when lagged effects as well as the effect of unobserved time-invariant confounders are stable over time, both the RI-CLPM and the DPM can recover the true effect. This characteristic of both models was replicated by the discussed simulations. However, the extent to which this is a reasonable causal model is debatable. For many phenomena, it may be more realistic to assume effects of confounders to vary over time. Furthermore, when times between measurements are not equal, time-stable lagged effect cannot be assumed either.

It is therefore necessary to expand the toolbox of social science researchers by investigating the use of causal inference methods in cross-lagged panel models, which will be covered by the thesis. The presented results show the necessity of such techniques and highlight the importance of acting with cause when attempting to assess cause and effect.

## References


Brown, D. W., Greene, T. J., Swartz, M. D., Wilkinson, A. V., & DeSantis, S. M. (2021). Propensity score stratification methods for continuous treatments. *Statistics in Medicine*, *40*(5), 1189–1203. https://doi.org/10.1002/sim.8835

Haber, N. A., Wieten, S. E., Rohrer, J. M., Arah, O. A., Tennant, P. W. G., Stuart, E. A., Murray, E. J., Pilleron, S., Lam, S. T., Riederer, E., Howcutt, S. J., Simmons, A. E., Leyrat, C., Schoenegger, P., Booman, A., Dufour, M.-S. K., O'Donoghue, A. L., Baglini, R., Do, S., … Fox, M. P. (2022). Causal and Associational Language in Observational Health Research: A Systematic Evaluation. *American Journal of Epidemiology*, *191*(12), 2084–2097. https://doi.org/10.1093/aje/kwac137

Hamaker, E. L. (2005). Conditions for the Equivalence of the Autoregressive Latent Trajectory Model and a Latent Growth Curve Model With Autoregressive Disturbances. *Sociological Methods & Research*, *33*(3), 404–416. https://doi.org/10.1177/0049124104270220

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, *38*(11), 2074–2102. https://doi.org/10.1002/sim.8086

Mulder, J. D. (2023). Power Analysis for the Random Intercept Cross-Lagged Panel Model Using the powRICLPM R-Package. *Structural Equation Modeling: A Multidisciplinary Journal*, *30*(4), 645–658. https://doi.org/10.1080/10705511.2022.2122467

Mulder, J. D., & Hamaker, E. L. (2021). Three Extensions of the Random Intercept Cross-Lagged Panel Model. *Structural Equation Modeling: A Multidisciplinary Jour-

*nal*, *28*(4), 638–648. https://doi.org/10.1080/10705511.2020.1784738

Murayama, K., & Gfrörer, T. (2022). *Thinking clearly about time-invariant confounders in cross-lagged panel models: A guide for choosing a statistical model from a causal inference perspective* [Preprint]. PsyArXiv.

Pearl, J., & Mackenzie, D. (2019). *The Book of Why: The New Science of Cause and Effect*. Penguin Books UK.

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Rosseel, Y. (2012). {Lavaan}: An {R} Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657. https://doi.org/10.1037/met0000210

Vansteelandt, S., & Daniel, R. m. (2014). On regression adjustment for the propensity score. *Statistics in Medicine*, *33*(23), 4053–4072. https://doi.org/10.1002/sim.6207