

Inferences and Effect Sizes for Direct, Indirect, and Total Effects in Continuous-Time Mediation Models

Ivan Jacob Agaloos Pesigan¹, Michael A. Russell^{1, 2}, and Sy-Miin Chow³

¹Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University

²Department of Biobehavioral Health, The Pennsylvania State University

³Department of Human Development and Family Studies, The Pennsylvania State University

Mediation modeling using longitudinal data is an exciting field that captures the interrelations in dynamic changes, such as mediated changes, over time. Even though discrete-time vector autoregressive (DT-VAR) approaches are commonly used to estimate indirect effects in longitudinal data, they have known limitations due to the dependency of inferential results on the time intervals between successive occasions and the assumption of regular spacing between measurements. Continuous-time vector autoregressive (CT-VAR) models have been proposed as an alternative to address these issues. Previous work in the area (e.g., Deboeck & Preacher, 2015; Ryan & Hamaker, 2021) has shown how the direct, indirect, and total effects, for a range of time interval values, can be calculated using parameters estimated from CT-VAR models for causal inferential purposes. However, both standardized effects size measures and methods for calculating the uncertainty around the direct, indirect, and total effects in continuous-time mediation have yet to be explored. Drawing from the mediation model literature, we present and compare results using the delta, Monte Carlo (MC), and parametric bootstrap (PB) methods to calculate standard errors (SEs) and confidence intervals (CIs) for the direct, indirect, and total effects in continuous-time mediation for inferential purposes. Options to automate these inferential procedures and facilitate interpretations are available in the cTMed R package.

Translational Abstract

Understanding how one variable influences another through an intermediate process is key in fields ranging from public health to behavioral science. Our study advances mediation modeling by using continuous-time models that capture how these relationships evolve dynamically—even when data are collected at irregular intervals. We introduce standardized effect size measures that adjust for differences in variable scales, enabling researchers to compare results across studies more easily. Additionally, we develop and compare methods, using delta, Monte Carlo, and parametric bootstrap techniques, to quantify the uncertainty in these mediation effects through confidence intervals and regions of significance. By implementing these methods in the open-source R package, cTMed, our work provides practical tools for researchers and practitioners, fostering better study designs, more accurate interpretations, and ultimately, more informed decision-making in applied research settings.

Supplementary materials: <https://raw.githubusercontent.com/jeksterslab/manCTMed/refs/heads/latex/manCTMed-supplement.pdf>.

Keywords: mediation, effect size, continuous-time, delta method, Monte Carlo method, parametric bootstrap, regions of significance, R package, cTMed

Mediation is a widely adopted model across the social, behavioral, and health sciences. A recent Google Scholar search, conducted as of May 8, 2024, using search terms "indirect effect OR mediation AND social science OR behavioral science OR health science", yielded over five million entries. Even though most mediation work is formulated predominantly as cross-sectional models (Figure 1a), there is a growing trend toward employing longitudinal designs (e.g., Figures 1b, 1c, and 1d) to address the critical issue of temporal ordering needed to facilitate (causal) interpretations of mediation effects (Cole

& Maxwell, 2003; Georgeson et al., 2023; Maxwell & Cole, 2007; Maxwell et al., 2011; O'Laughlin et al., 2018; Rogosa, 1979). Mediation models help us uncover the intermediate steps through which an independent variable X (such as an intervention) influences a dependent variable Y (such as an outcome) through one or more mediator variables M (e.g., Baron & Kenny, 1986; MacKinnon, 2008). By identifying these mediating variables, we gain insights into the direct and indirect causal pathways that connect interventions to their effects (Fairchild & MacKinnon, 2014; MacKinnon,

1994; Mackinnon & Dwyer, 1993; O'Rourke & MacKinnon, 2018, 2019).

This manuscript aims to provide methods to quantify and visualize the uncertainty associated with point estimates of direct, indirect, and total effects as they change across a range of time interval values in continuous time (CT) using the continuous-time vector autoregressive (CT-VAR) framework. We do this with repeated measurements over time as a more substantial basis for causal inference¹ in contrast to cross-sectional designs (Gollob & Reichardt, 1987, 1991). Furthermore, we take the CT framework as previous research (e.g., Deboeck & Preacher, 2015; Ryan & Hamaker, 2021) has shown its benefits, particularly in addressing the time interval dependency problem, that is, direct, indirect, and total effect sizes change as a function of the time interval used to measure the variables.

Our novel contributions in this paper include the following. First, we introduce standardized direct, indirect, and total effects as scale-independent effect size measures derived from model-implied steady-state variances. Like standardized regression coefficients, these metrics adjust for differences in measurement scales, making them interpretable and comparable across studies. The standardized indirect effect, expressed in standard deviation units, provides a consistent and meaningful index of mediation strength—especially valuable in the social, behavioral, and health sciences, where measurement units often vary widely (M. W.-L. Cheung, 2009a; Lachowicz et al., 2018; Preacher & Kelley, 2011). These metrics offer several advantages: they enhance inter-

pretability for diverse audiences, support cross-study comparisons and meta-analytic synthesis, and perform well in smaller samples, provided that appropriate confidence intervals are used to account for estimation uncertainty (M. W.-L. Cheung, 2009a; Lachowicz et al., 2018; Preacher & Kelley, 2011).

Second, we extend the inferential framework by developing methods to quantify the uncertainty associated with these effects. Specifically, we apply and compare the delta method (Casella & Berger, 2002; Oehlert, 1992; Ver Hoef, 2012), the Monte Carlo (MC) method (MacKinnon et al., 2004; Pesigan & Cheung, 2023; Preacher & Selig, 2012), and the parametric bootstrap (PB) method (Davison & Hinkley, 1997; Efron & Tibshirani, 1993; Preacher & Selig, 2012) to generate standard errors (SEs) and confidence intervals (CIs) for both the unstandardized and standardized effects across a range of time intervals. Furthermore, we propose using the delta, MC, and PB methods to generate regions of significance, which help depict and visualize the range of time interval values over which the direct, indirect, and total effects are statistically different from zero. It is important to note that these regions of significance are influenced by factors that affect statistical power, such as sample size and the reliability of measurement scales. Additionally, these regions represent confidence intervals, not credible intervals, and should be interpreted accordingly. Lastly, the methods we developed and visualization tools are available in a free and open-source package *cTMed*², which is implemented in R (R Core Team, 2025).

This manuscript is organized into the following sections: (1) introduction of CT mediation and the need to quantify the uncertainty associated with mediation effects; (2) to illustrate, using a benchmark example from the CT mediation literature, both the development of standardized effect size measures and the utility of performing inferences on mediation effects in CT using CIs obtained from the delta, MC, and PB methods, and regions of significance across a range of time intervals; and (3) two Monte Carlo simulation studies examining the respective finite-sample performance of these CIs. We close with a summary of our novel contributions, limitations of the current work, and future directions.

Mediation Using Longitudinal Data

This section discusses the rationale for moving from DT to CT mediation. We provide an overview of the CT mod-

¹Ivan Jacob Agaloos Pesigan; ²Michael A. Russell; ³Sy Miin Chow.

This research was made possible by the Prevention and Methodology Training Program (PAMT) funded by a T32 training grant (T32 DA017629 MPIs: J. Maggs & S. Lanza) from the National Institute on Drug Abuse (NIDA), the National Institutes of Health Intensive Longitudinal Health Behavior Cooperative Agreement Program U24AA027684, National Science Foundation grants DUE-2417294, the National Center for Advancing Translational Sciences under UL1TR002014-06, and the National Institute of Diabetes, Digestive & Kidney Diseases under U01DK135126.

Computations for this research were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer.

Some of the data and ideas in the manuscript were presented at the Society for Prevention Research Conference in May 2024 and the International Meeting of Psychometric Society in July 2024.

Correspondence concerning this article should be addressed to Ivan Jacob Agaloos Pesigan, Edna Bennett Pierce Prevention Research Center, College of Health and Human Development, The Pennsylvania State University, 320 Biobehavioral Health Building, University Park, PA 16802 or by email (ijapesigan@psu.edu).

This is an Author-Accepted Manuscript (AAM) accepted for publication in *Psychological Methods* on 11 June 2025.

¹For observational data, we limit our treatment of causality following the definition provided by Granger (1969), that is, when one variable, denoted as X , Granger-causes another variable, denoted as Y , it implies that past values of X contain information that helps predict future values of Y beyond what can be predicted using only past values of Y itself.

²See <https://CRAN.R-project.org/package=cTMed>. The specific version used is 1.0.6.

eling framework, namely the linear stochastic differential equation model. Although applications of the CT framework for mediation models exist in the literature (e.g., Deboeck & Preacher, 2015; Ryan & Hamaker, 2021), methods to quantify the uncertainty associated with mediation effects for inferential purposes have not yet been developed. The last part of this section outlines how we address this research gap.

From Discrete-Time to Continuous-Time Mediation

Mediation and causality have always been linked. From the *causal steps* approach (e.g., Baron & Kenny, 1986; James & Brett, 1984; Judd & Kenny, 1981) to the longitudinal framework (e.g., Cole & Maxwell, 2003; Gollob & Reichardt, 1987, 1991; Maxwell & Cole, 2007; Maxwell et al., 2011), the idea is that causality operates not only directly ($X \rightarrow Y$), but also indirectly ($X \rightarrow M \rightarrow Y$). Gollob and Reichardt (1991) proposed three principles of causality, namely: (1) “causes take time to exert their effects and, therefore, causal variables must occur before outcome variables”; (2) “variables can have effects on themselves”; (3) “size of the effect typically varies with the length of time lag” (p. 245–247).

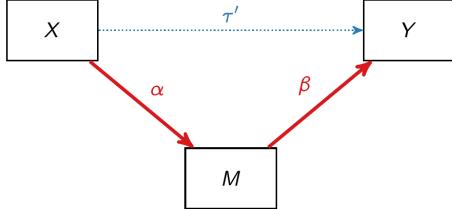
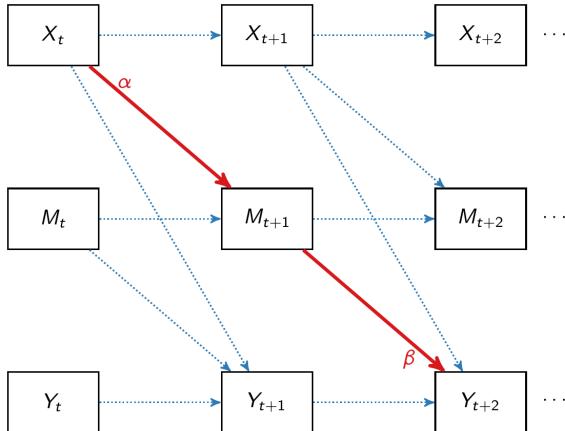
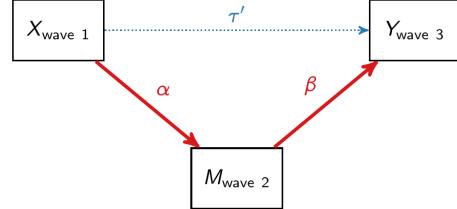
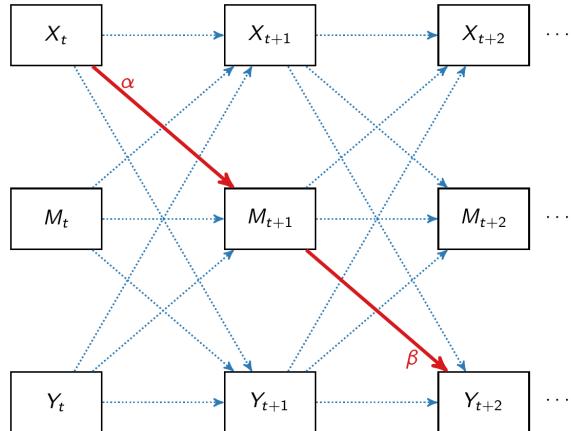
Principles 1 and 2 emphasize that cross-sectional mediation (Figure 1a), in itself, does not capture the direct and indirect causal links between variables X , M , and Y . While three-wave longitudinal models (Figure 1b) allow effects to unfold over time, they still fall short of satisfying the second principle. Instead, we should adopt an autoregressive cross-lagged mediation model (Figures 1c and 1d) that considers the previous variable values to predict their current values. The term autoregressive (AR) refers to allowing the variables to have effects on themselves over time, and the term cross-lagged or cross-regressive (CR) refers to using previous values of other variables to predict the current values of a variable. The umbrella term used for these types of models is the discrete-time vector autoregressive model (DT-VAR). The term *vector* autoregressive (VAR) means that there is more than one variable in the system. The term discrete-time (DT) means that we observe the variables at equally spaced time intervals (de Haan-Rietdijk et al., 2017). The DT-VAR framework encompasses a broad class of models, including cross-lagged panel models commonly applied in panel designs with relatively few time points (e.g., fewer than 5; see Finkel, 1995; Rogosa, 1980; Usami, 2020; Usami et al., 2019; Zyphur, Allison, et al., 2019; Zyphur, Voelkle, et al., 2019), as well as intensive longitudinal designs characterized by many repeated observations (e.g., 50 or more; see Hamilton, 1994; Lütkepohl, 2005). While these numbers are not strict thresholds, we include them to illustrate the range of study designs that can be modeled under the DT-VAR framework. In this manuscript, we treat DT-VAR as a general modeling framework that accommodates both panel and time series designs

as special cases by viewing individuals as replications of the same underlying dynamic process.

Principle 3 underscores the time interval dependency, that is, in DT-VAR, effect sizes change as a function of the time interval used to measure the variables (e.g., Deboeck & Preacher, 2015; Ryan & Hamaker, 2021; Wang & Zhang, 2020). This means that, given the same underlying causal process, the time elapsed (e.g., daily, monthly, yearly) between a “cause” (i.e., the independent variable) and its effect (i.e., the dependent variable) influences the sizes of the effects we observe.

In a cross-sectional mediation, if X changes by one unit, then Y changes by τ' units, holding M constant. This is the direct effect of X on Y . If X changes by one unit, then M changes by α units as a result. Similarly, if M changes by one unit, then Y changes by β units as a result. The resulting change of α units on M produced by a unit change in X would then produce a corresponding change of β units on Y . Thus, Y changes by $\alpha\beta$ units through M as X changes by one unit. This is the indirect effect. The total effect is the sum of direct and indirect effects. In DT-VAR, this idea of decomposing effects also holds. However, we are tracing effects from previous values of the variables to the next. The indirect effect of X at time t on Y at time $t + 2$ is the product of the path from X at time t to M at time $t + 1$ and the path from M at time $t + 1$ to Y at time $t + 2$. While many mediation models still rely on cross-sectional data, there is a growing trend toward employing DT-VAR models, as evidenced by the decrease in cross-sectional studies from the systematic review by Maxwell and Cole (2007) to that reported in O’Laughlin et al. (2018). However, the indirect effects and all the other effects we observe in DT-VAR change on account of the time elapsed between the observed variables. Thus, temporal precedence, one of the properties that make causal interpretations plausible, also generates new challenges, especially in cross-validating mediation results across studies and designs. A change in the modeling framework from DT to CT can address this issue.

The challenges associated with DT-VAR can be addressed by adopting a CT-VAR modeling framework. CT-VAR parameters are defined at any time interval; as such we can obtain the direct, indirect, and total effects for any time interval. From this framework, we can explore how effects change as a function of time intervals. Additionally, unlike DT-VAR, which assumes equally spaced data points, CT-VAR accommodates unequally spaced data—a common occurrence in intensive longitudinal studies. These models allow us to capture the evolving dynamics of parameters across a range of time intervals (Boker, 2002; Oud & Delsing, 2010; van Montfort et al., 2018; Voelkle et al., 2012). However, embracing a CT perspective necessitates a conceptual shift in the definition of “total” effect (Aalen et al., 2012, 2016; De-

Figure 1*Cross-Sectional and Longitudinal Mediation Models*(a) *Cross-sectional mediation model.*(c) *Unidirectional DT-VAR.*(b) *Three-wave longitudinal mediation model.*(d) *Bidirectional DT-VAR.*

Note: The indirect effect of M on the path from X to Y is the product of α and β . For cross-sectional mediation (1a), variables are typically measured at the same time point. The three-wave longitudinal mediation (1b) is an improvement to cross-sectional mediation where variables are measured at three distinct time points. Waves 1, 2, and 3 may or may not be equally spaced. DT-VAR mediation goes further by allowing variables to have effects on themselves. Here t , $t + 1$, and $t + 2$ typically represent equally spaced time points. The dots indicate that the model can accommodate more time points. We distinguish between the unidirectional model (1c), where paths from M to X , Y to M and Y to X are constrained to zero and the bidirectional model (1d), where all paths from the previous time point to the current are estimated.

boeck & Preacher, 2015; Ryan & Hamaker, 2021) using CT terminologies, which we review briefly next.

CT Modeling Using Linear Stochastic Differential Equations (SDE)

In our modeling approach, we consider processes that evolve continuously over time. This means that the variables can assume values at any given moment. To capture this behavior, we employ linear stochastic differential equations (SDEs). Specifically, we focus on a special type of SDE known as the Ornstein–Uhlenbeck model (Chow et al., 2023; Oravecz et al., 2011; Uhlenbeck & Ornstein, 1930).

Similar to structural equation modeling, we can break down our model into two key components: the measurement component and the structural component (Chow et al., 2010). The measurement model describes the relationship between observed variables and latent (unobserved) variables. Specif-

ically, consider the following equation:

$$\mathbf{y}_{i,t_m} = \boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\eta}_{i,t_m} + \boldsymbol{\varepsilon}_{i,t_m}, \quad \boldsymbol{\varepsilon}_{i,t_m} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Theta}) \quad (1)$$

where \mathbf{y}_{i,t_m} represents a vector of observed random variables, $\boldsymbol{\eta}_{i,t_m}$ a vector of latent random variables, and $\boldsymbol{\varepsilon}_{i,t_m}$ a vector of normally distributed random measurement errors, for time t at measurement occasion m and individual i . $\boldsymbol{\nu}$ denotes a vector of intercepts, $\boldsymbol{\Lambda}$ a matrix of factor loadings, and $\boldsymbol{\Theta}$ the covariance matrix of $\boldsymbol{\varepsilon}$. In the context of a mediation model with three variables where the effect of the variable X on the variable Y is mediated by the variable M , we define \mathbf{y} and $\boldsymbol{\eta}$ as follows:

$$\mathbf{y} = \begin{pmatrix} X \\ M \\ Y \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\eta} = \begin{pmatrix} \eta_X \\ \eta_M \\ \eta_Y \end{pmatrix}. \quad (2)$$

The structural model for $\boldsymbol{\eta}$ in a linear SDE framework focuses on delineating how $\boldsymbol{\eta}$ changes over an infinitely small

time interval Δt , as Δt approaches 0. Using the Ornstein–Uhlenbeck model as the representation for change, for example, the structural model is commonly referred to as the *dynamic* model, and it is given by

$$d\eta_{i,t} = \Phi(\eta_{i,t} - \mu) dt + \Sigma^{\frac{1}{2}} dW_{i,t} \quad (3)$$

where μ is the equilibrium or mean value, Φ is the drift matrix, Σ is the matrix of volatility or randomness in the process (also known as the diffusion matrix), and dW is a Wiener process or Brownian motion, which represents random fluctuations. The Ornstein–Uhlenbeck process centers around a central tendency denoted by μ . When the process deviates from this mean, it naturally gravitates back toward it when the stability condition is satisfied, namely, the real parts of all eigenvalues of the matrix Φ are less than zero (Ryan et al., 2018). We can think of μ as an *attractor* that pulls the process back to its typical value. When we specify the model such that the latent state variables have a stable mean vector μ consisting of zeros, Equation 3 simplifies to

$$d\eta_{i,t} = (\Phi\eta_{i,t}) dt + \Sigma^{\frac{1}{2}} dW_{i,t}. \quad (4)$$

Elements of the matrix Φ control the speed at which each of the processes reverts to its equilibrium value, μ . The diagonal elements in Φ , referred to as the auto effects, govern how quickly the variable returns to its mean after being perturbed in the absence of influences from other processes. The off-diagonal elements of Φ represent associations between variables and their rates of change. A negative cross effect indicates that if one variable takes on a positive value, there is a resulting decrease in another variable's value. The magnitude of this effect is determined by the absolute value. The diagonal elements of the matrix Σ determine the magnitude of volatility around the central tendency μ . Larger values indicate greater fluctuations or noise in the process, while smaller values result in smoother trajectories. Essentially, Σ quantifies the degree of randomness or shock-induced variability. This is also referred to as dynamic error. The off-diagonal elements of Σ represent correlations between random shocks across variables in the system. In summary, under the stability condition, the Ornstein–Uhlenbeck process is a mean-reverting stochastic process. Over time, it drifts toward and stays at its equilibrium value μ unless unmeasured shocks (the magnitudes of which are captured by Σ) push it away from the equilibrium, in which case it then reverts to the mean with a rate determined by Φ . For the context of this manuscript, we refer to the Ornstein–Uhlenbeck model as the first-order CT-VAR.

The drift matrix Φ holds particular significance in the context of mediation models. It serves as a conduit for understanding the associations between the states of the variables under investigation. Notably, within a trivariate model, involving η_X , η_M , and η_Y , numerous paths from one variable

to another can be explored. This comprehensive connectivity arises because we can construct the model such that all variables are interconnected. Consequently, beyond the classic mediation model $\eta_X \rightarrow \eta_M \rightarrow \eta_Y$, we can also investigate alternative mediation pathways, such as $\eta_Y \rightarrow \eta_M \rightarrow \eta_X$, using the same data set. For our current focus, we will delve into the specific $\eta_X \rightarrow \eta_M \rightarrow \eta_Y$ path. In the subsequent section, we will explore the relationship between CT-VAR and DT-VAR models, emphasizing the pivotal role played by the drift matrix Φ within the $\eta_X \rightarrow \eta_M \rightarrow \eta_Y$ mediation model.

Link Between CT-VAR and DT-VAR

To understand the connection between CT-VAR and DT-VAR models, we need to examine the state space model parameterization of the CT-VAR model which is the DT solution in integral form given by

$$\eta_{i,t_{m_i}} = \alpha_{\Delta t_{m_i}} + \beta_{\Delta t_{m_i}} \eta_{i,t_{m_i-1}} + \zeta_{i,t_{m_i}}, \quad \zeta_{i,t_{m_i}} \sim \mathcal{N}(\mathbf{0}, \Psi_{\Delta t_{m_i}}) \quad (5)$$

where $\eta_{i,t_{m_i}}$ and $\eta_{i,t_{m_i-1}}$ represent the vector of latent variables in the current and previous measurement occasion m while $\zeta_{i,t_{m_i}}$ is the normally distributed error term in the current measurement occasion. The model parameter α denotes a vector of intercepts, β a matrix of coefficients describing the impact of the lagged values $\eta_{i,t_{m_i-1}}$ on the current value η_{i,t_m} (AR coefficients on the diagonal and CR coefficients on the off-diagonal), and Ψ the covariance matrix of ζ . Notice that the parameters have subscript Δt , representing that their values are dependent on the time interval, that is, given the same continuous process, their values change depending on the time interval. Note that, here, we assume stationarity, particularly weak stationarity, that is, the time series has a constant mean, and the covariance remains invariant over time³.

In the context of CT-VAR models, the underlying process can take values at any arbitrary time point. However, our observations occur discretely. This means that while the process itself operates continuously, we only observe it at specific discrete time instances. Let's compare this to DT-VAR, where observations are made at specific discrete time intervals (usually positive integers). In a DT-VAR example, suppose we collect data every hour from 7 AM to 10 AM for four measurement occasions. We make discrete observations in time t , scaled by the hour, on specific measurement occasions $m = \{1, \dots, 4\}$. These occasions correspond to time points: $t_1 = 7$, $t_2 = 8$, $t_3 = 9$, and $t_4 = 10$. The time interval from one measurement occasion to the next is constant and equal to 1 hour. In the CT-VAR model, observations can occur at any time between 7 AM and 10 AM for the same four measurement occasions. Specifically, the time

³While weak stationarity is a common assumption in CT-VAR, we acknowledge that it may not always hold in empirical panel data. We adopted this assumption to simplify the simulation design and to remain consistent with prevailing modeling practices.

points are given by: $t_1 = 7.166667$ (equivalent to 7:10 AM), $t_2 = 8.283333$ (equivalent to 8:17 AM), $t_3 = 9.45$ (equivalent to 9:27 AM), and $t_4 = 10$ (equivalent to 10:00 AM). Notably, the time interval from one measurement occasion to the next varies, allowing for irregularly spaced observations. We denote this time interval as Δt , representing the difference between the current measurement occasion m and the previous measurement occasion $m - 1$. From here we can see that if Δt is constant as in DT-VAR, the model is simply a CT-VAR model where Δt is constant. As such, the DT-VAR model is a special case of the CT-VAR model.

Let us look at the link between the CT-VAR and the DT solution parameters. Whereas the parameter Φ describes the moment-to-moment association of one variable on the rate of change in another, the parameter β is a matrix of lagged parameters, that is, the effect of the variable on a previous measurement occasion to the current measurement occasion, given a particular time interval Δt . The DT β can be obtained from the CT Φ using

$$\beta = \exp(\Delta t \Phi) \quad (6)$$

where $\exp(\cdot)$ represents the matrix exponential function.

In addition, we can convert the CT process noise covariance matrix Σ to the DT counterpart Ψ as:

$$\text{vec}(\Psi) = \mathbf{C} [\exp(\Delta t \mathbf{C}) - \mathbf{I}] \text{vec}(\Sigma), \quad \mathbf{C} = (\Phi \otimes \mathbf{I}) + (\mathbf{I} \otimes \Phi) \quad (7)$$

where \mathbf{I} is an identity matrix, $\text{vec}(\cdot)$ stacks the values from a matrix into a vector by column, and \otimes is the Kronecker product⁴.

In summary, we can see the DT-VAR parameters are related to CT-VAR parameters. In the special case where the time intervals across occasions are constant, the CT-VAR reduces to a DT-VAR. As Δt is part of the calculation of the DT parameters, their values change as a function of the time interval. For more information on CT-VAR models see Chow et al. (2023), Oravecz et al. (2011), Ryan et al. (2018), and Voelkle and Oud (2012).

Lagged Parameters β as Matrix of Total Effects

An important element of CT mediation modeling involves a conceptual shift to viewing lagged parameters β as total rather than direct effects. This view is supported by works such as Aalen et al. (2012, 2016), Deboeck and Preacher (2015), and Ryan and Hamaker (2021). This shift is particularly relevant in CT modeling and emphasizes understanding relationships as integrated impacts over time rather than discrete, isolated causations. One key rationale for this approach is the issue of time-interval dependency inherent in DT-VAR models. As noted by Gollob and Reichardt (1987) and expanded upon by Ryan and Hamaker (2021), lagged parameters in DT models depend on the specific time intervals between observations. This dependency complicates the

comparison of findings across studies with different sampling intervals. CT models address this challenge by providing lag-independent parameters, allowing for consistent interpretation over various time intervals.

Another critical reason for treating lagged parameters as total effects is the integrated nature of causality over time. Aalen et al. (2012, 2016) argued that causality in CT processes requires consideration that these processes exist and continue to evolve over time—even when they are not being measured. Consequently, direct effects between any two constructs should *exclude* indirect effects that unfold via other constructs between observed time points. This approach is particularly insightful in mediation models within autoregressive frameworks, as it accounts for the cumulative impact of mediators and outcomes over time, providing a more complete picture of causality. As an example, this perspective advocates that in Figures 1c and 1d, the direct effect of X_t on Y_{t+1} would include the effect of X_t on Y_{t+2} through X_{t+1} , but exclude the effect of M_t on X_{t+1} to filter out any implicit or explicit influence of M . As such, the whole matrix of paths linking X , Y , and M between time t , $t + 1$, and $t + 2$ specifies the total effects of all these variables on each other, including both direct and indirect effects.

The shift to total effects also aligns with a mechanistic understanding of processes, a concept emphasized by Aalen et al. (2012, 2016). The mechanistic perspective prioritizes understanding the dynamic behavior of entire systems rather than focusing solely on relationships at isolated, punctuated observed time points. This is valuable when studying mediation, as it offers a more holistic view of how relationships between variables collectively influence outcomes over time. By interpreting lagged parameters as total effects, researchers can better capture the complexity of the system and the pathways that contribute to the observed results. Additionally, CT models enable consistent interpretations across different time scales, avoiding the issue of time-interval dependency and making results more generalizable. Collectively, these advantages provide deeper insights into the mechanisms underlying observed effects, facilitating a richer understanding of how outcomes are produced.

Quantifying Uncertainty in Direct, Indirect, and Total Effects in Continuous-Time Mediation Models

The work of Deboeck and Preacher (2015) provided a starting point for structuring mediation processes in CT-VAR. However, their work focused on point estimates and did not provide explicit direction on how to calculate SEs and

⁴The notation presented here is mathematically equivalent to the notation presented in the works of Arnold (1974), Oud and Jansen (2000), and Voelkle and Oud (2012). We prefer this notation because it facilitates straightforward calculations in statistical packages that support standard linear algebra operations.

CIs. Complete inferences require the availability of reasonable SE estimates and ways of quantifying the uncertainty around the direct, indirect, and total effects in mediation in CT. Of particular interest in the mediation literature are inferences concerning indirect effects. Although a point estimate of the indirect effect can be obtained straightforwardly as a product of regression coefficients, the sampling distribution of the product of normally distributed random variables is not necessarily normal (Aroian, 1947; Craig, 1936), leading to complexities in computing p -values for the indirect effect. Because of this, CIs have been used to quantify uncertainty for the point estimates. Using the delta method, CIs that are asymptotically normally distributed have been proposed (Aroian, 1947; Goodman, 1960; Sobel, 1982, 1986, 1987).

Alternative approaches that take into account the asymmetry have also been proposed including: (1) The distribution of the product method, which analytically examines the sampling distribution of the product of normally distributed coefficients (MacKinnon et al., 2004, 2007); (2) Nonparametric bootstrapping (NB), involving resampling the original data and fitting the model multiple times to create a sampling distribution of the indirect effect (Bollen & Stine, 1990; S. F. Cheung et al., 2022; Preacher & Hayes, 2008; Shrout & Bolger, 2002); (3) Parametric bootstrapping (PB), involving simulating raw data using the estimated parameters and fitting the model similar to NB (Preacher & Selig, 2012); (4) Monte Carlo method (MC), which uses point estimates and SEs to form parametric distributions from which to draw values of regression coefficients to form a sampling distribution of the indirect effect (MacKinnon et al., 2004; Pesigan & Cheung, 2023; Preacher & Selig, 2012)⁵; (5) Profile likelihood-based approach that estimate CIs using the likelihood ratio test (M. W.-L. Cheung, 2009a, 2009b; S. F. Cheung & Pesigan, 2023; Pawitan, 2013; Pesigan & Cheung, 2020; Venzon & Moolgavkar, 1988). From these approaches, NB has become the preferred method (Bollen & Stine, 1990; Preacher & Hayes, 2008; Shrout & Bolger, 2002). However, the need to iterate the estimation process repeatedly for all the (usually thousands of) bootstrap data sets creates a considerable computational burden, especially for complex models (see Pesigan & Cheung, 2023, for the case of mediation models with missing data). MC is as good as NB approaches in many situations (Pesigan & Cheung, 2023; Tofghi & Kelley, 2020; Tofghi & MacKinnon, 2015). Because the empirical sampling distributions of parameter estimates are generated and not obtained via multiple model fitting, MC is much more computationally efficient than NB and PB.

In this manuscript, drawing from the mediation literature (e.g., MacKinnon et al., 2004; Pesigan & Cheung, 2023; Preacher & Selig, 2012; Sobel, 1986), we provide estimation procedures to complete the inferential process in CT-VAR mediation using the delta, MC, and PB methods. We chose the delta method because of its generality and ease

of implementation and MC as a more computationally efficient alternative to the bootstrap. We selected PB over the more commonly used NB because PB better preserves the temporal dependencies inherent in CT-VAR models. Traditional NB methods assume independent and identically distributed observations, an assumption clearly violated in time series contexts like CT-VAR where data are inherently auto-correlated. While various time-series-specific bootstrapping approaches exist—such as block bootstrapping or stationary bootstrapping (Kreiss & Lahiri, 2012; Kunsch, 1989; Lahiri, 2003; Politis, 2003; Politis & Romano, 1994; Shumway & Stoffer, 2017; Stoffer & Wall, 1991)—these methods require selecting tuning parameters (e.g., block length), rely on assumptions about stationarity, and often depend on model-specific implementations to replicate the dependency structure. In contrast, PB directly simulates data from the estimated CT-VAR model, ensuring that both the stochastic and dynamic structures of the data are preserved without the need for resampling decisions or simplifying assumptions. This approach is further motivated by the strong performance of PB under correctly specified discrete-time dynamic models (Zhang, 2018; Zhang & Browne, 2010), suggesting its utility as a benchmark in time-dependent mediation contexts. Although PB is more computationally intensive than the delta and MC methods, its ability to produce valid inferences in structured time-dependent data makes it a theoretically grounded and practically robust choice for comparison.

To facilitate implementation, we developed the **cTMed** R package, which provides routines for computing and visualizing unstandardized and standardized direct, indirect, and total effect CIs using delta, MC, and PB methods. **cTMed** goes beyond existing packages such as **dynr** (Ou et al., 2019), **OpenMx** (Hunter, 2017; Neale et al., 2015), **ctsem** (Driver & Voelkle, 2018; Driver et al., 2017), and **Mplus** (Muthén & Muthén, 2017). These alternative software programs provide general functionalities for users to obtain SEs and CIs through the delta method for any special cases of their broadest modeling frameworks, but require explicit specifications of the desired transformations and observed (possibly irregular) time intervals of the data, which can be unwieldy in high-dimensional networks. Instead, **cTMed** automates and provides readily available tools to calculate and visualize inferential results for the total, direct, and indirect effects (unstandardized and standardized) within CT mediation models through the delta-, MC-, and PB-based CIs. This intuitive, user-friendly interface automates key procedures for making inferences in CT mediation models, making the process comparable in accessibility to standard cross-sectional mediation models.

⁵Not to be confused with Bayesian posterior simulation methods.

Three Illustrations Using an Empirical Data Set

In this section, we present three illustrations that address specific challenges that arise in making inferences with CT mediation models based on one empirical data set. These illustrations serve to: (1) show how standardized direct, indirect, and total effects can be obtained from point estimates of CT mediation models, (2) develop and compare the performances of the delta method, MC method, and PB in deriving SEs and CIs for CT mediation models; and (3) demonstrate use of graphical tools to display and facilitate interpretations of mediation effects in CT models across different time scales.

Illustration 1: Obtaining Standardized Direct, Indirect, and Total Mediation Effects in CT-VAR Models

In this section, we illustrate how to compute direct, indirect, and total effects from the estimated drift matrix Φ in a three-variable CT-VAR model with bidirectional influences. We also examine how these effects change as a function of time intervals. In both cross-sectional and autoregressive models, mediation analysis involves decomposing the total effect into direct and indirect components. This approach helps us understand how the mediator variable M influences the relationship between X and Y . However, in the context of autoregressive models, some authors (e.g., Aalen et al., 2012, 2016; Deboeck & Preacher, 2015; Ryan & Hamaker, 2021) propose a conceptual shift when thinking about the lagged parameters β . Here, β is treated as the matrix of total effects rather than direct effects. This shift has implications for mediation in VAR models, particularly when decomposing the total effect into direct and indirect effects. In particular, we focus on the “difference approach,” which defines direct effects between any two implicated variables as those that involve these variables solely, while indirect effects represent the remaining influence after subtracting direct effects from total effects. This approach contrasts with the traditional method, where the indirect effect is obtained by multiplying path coefficients along the mediation pathway. In CT mediation models, the total, direct, and indirect effects are represented as sums of products of coefficients over time. Specifically, the indirect effect is derived by subtracting the direct effect (which holds mediators constant) from the total effect (which includes all possible pathways).

Example of Matrices of Direct, Indirect, and Total Effects

In this illustration, we used the CT-VAR drift matrix reported in Table 2 of Deboeck and Preacher (2015) based on the results reported in Grundy et al. (2007) to demonstrate the decomposition of direct and indirect effects from the total effect. The data set was generated using the `manCTMed::IllustrationGenData`⁶ function which is a wrapper for the `SimSSMOFixed` function from the

`simStateSpace`⁷ R package. The original research examined whether maternal knowledge mediated the association between mother-reported marital conflict and child-reported behavioral competence, including a reciprocal effect from behavioral competence to knowledge for 133 mother and child pairs. We extended the illustration by fitting a CT-VAR model using the `manCTMed::IllustrationFitDynr` function which is a wrapper around the `dynr`⁸ R package (Ou et al., 2019), estimating both measurement and dynamic structures. Additionally, we allowed for bidirectional paths by estimating all possible paths from one variable to another in the drift matrix Φ .

We labeled the independent variable marital conflict X , the mediator variable maternal knowledge M , and the dependent variable behavioral competence Y . Since, we are incorporating a measurement model, the latent variables η_X , η_M , and η_Y represent the latent variables in the model with X , M , and Y as observed indicators. Elements of the estimated CT-VAR drift matrix Φ are presented in Table 1. In the study, time t is scaled in years, and $\Delta t = 1$ is a time interval of one year between measurement occasions given in Figure 2a. Given Equation 6 we can calculate the matrix of lagged coefficients for a time interval $\Delta t = 1$ from the drift matrix Φ which is also presented in Table 1 as follows:

$$\begin{aligned} \boldsymbol{\beta}_{\Delta t(1)} &= \exp(\Phi) \\ &= \eta_{M_{t+1}} \begin{pmatrix} \eta_{X_t} & \eta_{M_t} & \eta_{X_t} \\ \beta_{1,1} & \beta_{1,2} & \beta_{1,3} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} \\ \beta_{3,1} & \beta_{3,2} & \beta_{3,3} \end{pmatrix}. \end{aligned} \quad (8)$$

Now we look at the case of a DT-VAR model from the same continuous process except that the time interval used is three years ($\Delta t = 3$) between measurement occasions. From the time-series literature (e.g., Hamilton, 1994) we can derive $\boldsymbol{\beta}_{\Delta t(3)}$ from $\boldsymbol{\beta}_{\Delta t(1)}$ using matrix powers. The key is to raise the matrix with the smaller time interval by the factor in which the time interval increased to the larger time interval. For example, the change from $\Delta t = 1$ to $\Delta t = 3$ is a factor of three, thus $\boldsymbol{\beta}_{\Delta t(1)}^3 = \boldsymbol{\beta}_{\Delta t(3)}$. Therefore, we can say that $\boldsymbol{\beta}_{\Delta t(3)} = \exp(3\Phi) = \boldsymbol{\beta}_{\Delta t(1)}^3$. Let us explore the relationship of the elements of $\boldsymbol{\beta}_{\Delta t(3)}$ as a function of the elements of $\boldsymbol{\beta}_{\Delta t(1)}$. In particular, we focus on the third row and first column in $\boldsymbol{\beta}_{\Delta t(3)}$ pertaining to the association of marital conflict (η_X) at

⁶`manCTMed` is an R package hosted in GitHub (<https://github.com/jeksterslab/manCTMed>) serving as a research compendium for this manuscript. It contains all the scripts and other supplementary materials used throughout this work.

⁷See <https://CRAN.R-project.org/package=simStateSpace>. The specific version used is 1.2.9.

⁸The development version available in <https://github.com/mhuntern1/dynr> particularly commit `0ac88f919494054938d6d71d02f6b95ab6a58515` was used.

Table 1
Point Estimates and 95% CIs of the CT-VAR Drift Matrix Φ and the DT-VAR ($\Delta t = 1$) Lagged Coefficients Matrix β for a Single Data Set Simulated from Grundy et al. (2007)

Parameter	CT-VAR			DT-VAR			DT-VAR			
	Estimate	2.5%		Estimate	2.5%		MC	PB PC		PB BC
		97.5%	97.5%		97.5%	97.5%		97.5%	97.5%	
Conflict to conflict ($X \rightarrow X$)	-0.116	-0.150	-0.082	0.890	0.861	0.919	0.862	0.919	0.861	0.918
Conflict to knowledge ($X \rightarrow M$)	-0.123	-0.159	-0.086	-0.111	-0.067	-0.110	-0.067	-0.112	-0.067	-0.067
Conflict to competence ($X \rightarrow Y$)	-0.088	-0.123	-0.053	-0.063	-0.085	-0.040	-0.084	-0.040	-0.086	-0.040
Knowledge to conflict ($M \rightarrow X$)	0.020	-0.088	0.128	0.013	-0.054	0.079	-0.053	0.079	-0.056	0.083
Knowledge to knowledge ($M \rightarrow M$)	-0.882	-1.010	-0.754	0.420	0.366	0.473	0.369	0.477	0.362	0.472
Knowledge to competence ($M \rightarrow Y$)	0.060	-0.055	0.175	0.027	-0.025	0.080	-0.027	0.078	-0.022	0.079
Competence to conflict ($Y \rightarrow X$)	-0.007	-0.112	0.098	-0.002	-0.067	0.062	-0.066	0.063	-0.065	0.061
Competence to knowledge ($Y \rightarrow M$)	0.474	0.357	0.590	0.220	0.172	0.269	0.171	0.268	0.174	0.268
Competence to competence ($Y \rightarrow Y$)	-0.064	-0.784	-0.543	0.522	0.466	0.577	0.470	0.580	0.463	0.576

Note. Estimates for the CT-VAR are calculated using the dcmr package. Estimates for the DT-VAR are calculated using the DeltaBeta (delta method), MCbeta (MC), and BootBeta (PB) functions from the cTfied package.

time t and behavioral competence (η_Y) at time $t + 3$. The dashed path in Panel 2a in Figure 2 represents the total effect of marital conflict (η_X) at time t on behavioral competence (η_Y) at time $t + 3$. This total effect is decomposed as a function of the lagged parameters in $\beta_{\Delta t(1)}$. By raising Equation 8 to the third power and subsetting the third row and the first column of the resulting matrix we get

$$\begin{aligned} \text{Total} &= \beta_{\Delta t(3)} [3, 1] \\ &= \beta_{1,1}\beta_{1,1}\beta_{3,1} + \beta_{2,1}\beta_{1,2}\beta_{3,1} + \beta_{3,1}\beta_{1,3}\beta_{3,1} \\ &\quad + \beta_{1,1}\beta_{2,1}\beta_{3,2} + \beta_{2,1}\beta_{2,2}\beta_{3,2} + \beta_{3,1}\beta_{2,3}\beta_{3,2} \\ &\quad + \beta_{1,1}\beta_{3,1}\beta_{3,3} + \beta_{2,1}\beta_{3,2}\beta_{3,3} + \beta_{3,1}\beta_{3,3}\beta_{3,3}, \end{aligned} \quad (9)$$

where each term represents distinct paths from marital conflict (η_X) at time t to behavioral competence (η_Y) at time $t + 3$ graphically represented by Panels 2b to 2j in Figure 2. Note that solid paths are paths where maternal knowledge (η_M) is traversed. We define these solid paths as indirect effects of marital conflict (η_X) at time t on behavioral competence (η_Y) at time $t + 3$ through maternal knowledge (η_M). The sum of these solid paths constitutes the overall indirect effect. The dashed paths are the direct effects and their sum constitutes the overall direct effect of marital conflict (η_X) at time t on behavioral competence (η_Y) at time $t + 3$. The sum of the direct and indirect effects represents the total effect of marital conflict (η_X) at time t on behavioral competence (η_Y) at time $t + 3$.

Now we divide any arbitrary time interval Δt into infinitely many smaller parts and insert these infinitely many latent variables between values at time t and time $t + \Delta t$. The idea in the CT-VAR approach is that the lagged parameters from the DT integral solution represent the sum of all the effects from the infinitely many smaller parts in between time t and time $t + \Delta t$. The decomposition we discussed earlier still applies. The total effect is straightforward as it is simply the result of exponentiating the product of the drift matrix and the time interval. Using the estimated drift matrix presented in Table 1 and a time interval $\Delta t = 3$, the total effect matrix is given by

$$\begin{aligned} \text{Total} &= \exp(\Delta t \Phi) \\ &= \exp \left[3 \begin{pmatrix} -0.116 & 0.020 & -0.007 \\ -0.123 & -0.882 & 0.474 \\ -0.088 & 0.060 & -0.664 \end{pmatrix} \right] \\ &= \begin{pmatrix} 0.703 & 0.017 & 0.002 \\ -0.145 & 0.080 & 0.148 \\ -0.100 & 0.017 & 0.151 \end{pmatrix}. \end{aligned} \quad (10)$$

Following Ryan and Hamaker (2021), the matrix of direct effects is obtained by removing the effects in the drift matrix Φ that originates from and/or goes to the mediator (or mediators) from Φ before converting it into its DT counterpart, β , for a given time interval, Δt . We do this by pre- and post-multiplying the drift matrix Φ with a diagonal matrix \mathbf{D}_m

where the i^{th} and j^{th} index pertaining to the mediators are set to zero. In our current example where maternal knowledge (η_M) is treated as the mediator variable, the direct effect matrix is given by

$$\begin{aligned} \text{Direct} &= \exp(\Delta t \mathbf{D}_m \Phi \mathbf{D}_m) \\ &= \exp \left[3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -0.116 & 0.020 & -0.007 \\ -0.123 & -0.882 & 0.474 \\ -0.088 & 0.060 & -0.664 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \\ &= \begin{pmatrix} 0.708 & 0.000 & -0.007 \\ 0.000 & 1.000 & 0.000 \\ -0.091 & 0.000 & 0.137 \end{pmatrix}. \end{aligned} \quad (11)$$

In other words, the direct effects may be conceptualized as effects that are due only to direct linkages (i.e., arrows) between the “independent variable” and “dependent variable.” The indirect effect, in contrast, can be understood as all the influences that involve paths that go through the mediators, which can now be obtained by subtracting the direct effect matrix from the total effect matrix as:

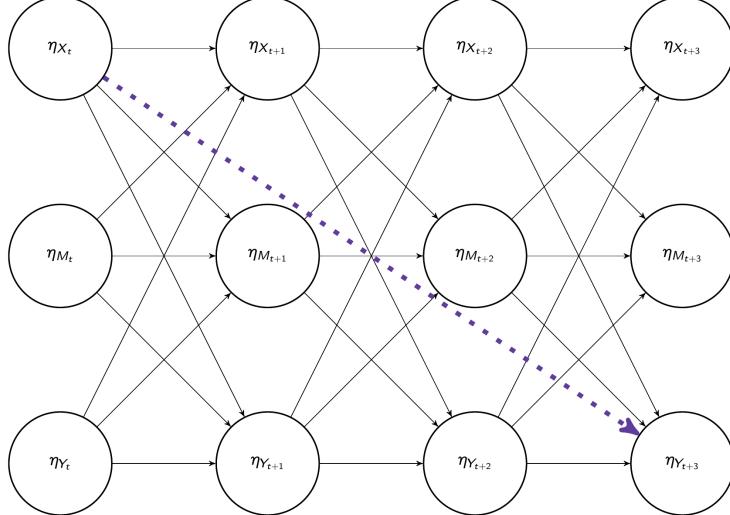
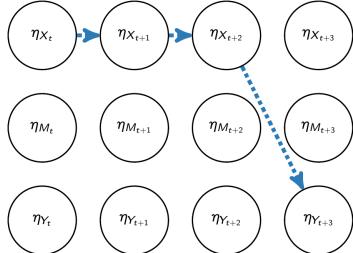
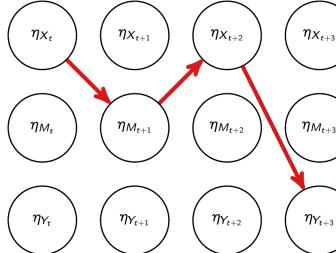
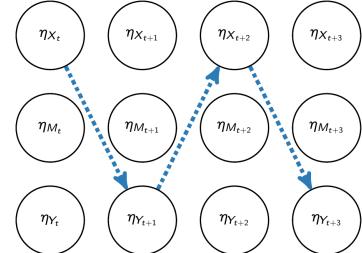
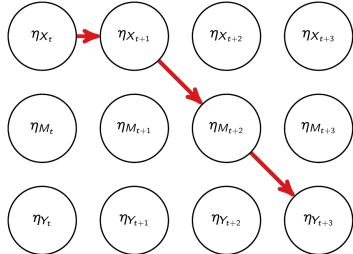
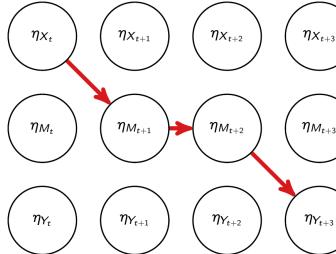
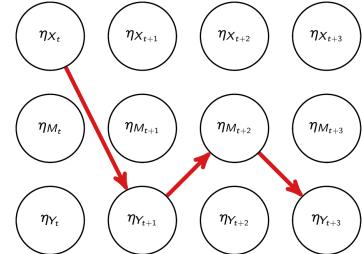
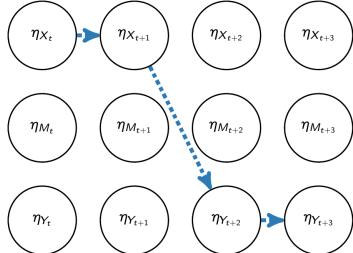
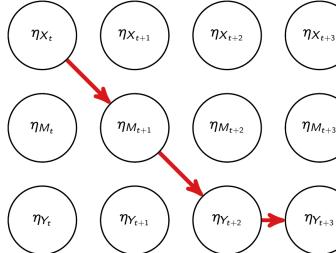
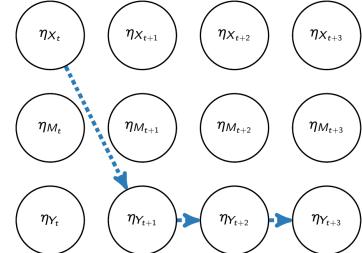
$$\begin{aligned} \text{Indirect} &= \text{Total} - \text{Direct} \\ &= \exp(\Delta t \Phi) - \exp(\Delta t \mathbf{D}_m \Phi \mathbf{D}_m) \\ &= \begin{pmatrix} -0.005 & 0.017 & 0.009 \\ -0.145 & -0.920 & 0.148 \\ -0.009 & 0.017 & 0.014 \end{pmatrix}. \end{aligned} \quad (12)$$

We subset the i^{th} row and j^{th} column of the resulting matrices for the direct, indirect, and total effects of the mediator variable/s on the path from the j^{th} column variable (1 for marital conflict in our case) to the i^{th} row variable (3 for behavioral competence in our case). The direct, indirect, and total effects of marital conflict (η_X) on behavioral competence (η_Y) through maternal knowledge (η_M), for a time interval of 3, are -0.091 , -0.009 , and -0.100 , respectively⁹.

If we constrain the upper diagonal elements of the drift matrix to zeroes, the matrix of lagged coefficients will also have zeroes in the upper diagonals. We described this previously as the unidirectional DT-VAR model shown in Figure 1c. The approach described here—following Ryan and Hamaker (2021)—is equivalent to the method proposed by Deboeck and Preacher (2015) when these constraints are in place. The key advantage of the current approach is its greater generality, as it also accommodates the bidirectional DT-VAR model depicted in Figure 1d when the upper diagonal elements of the drift matrix are freely estimated.

The sum of all the effects that pass from marital conflict to behavioral competence, which we call the total effect, is -0.100 . This is the accumulated effect up until the third year (given $\Delta t = 3$ and time t scaled in terms of years). We can decompose this total effect in terms of the effect that passes

⁹The Med function in the cTMed package calculates the direct, indirect, and total effects for a specific time interval or a range of time intervals. A plot method for the Med function can be used to visualize the effects as a function of time intervals.

Figure 2*Decomposing the Direct and Indirect Effects from the Total Effect*(a) DT-VAR with a time interval of one. The dashed path represents the total effect of X on Y at a time interval of three.(b) $\beta_{1,1}\beta_{1,1}\beta_{3,1}$ (c) $\beta_{2,1}\beta_{1,2}\beta_{3,1}$ (d) $\beta_{3,1}\beta_{1,3}\beta_{3,1}$ (e) $\beta_{1,1}\beta_{2,1}\beta_{3,2}$ (f) $\beta_{2,1}\beta_{2,2}\beta_{3,2}$ (g) $\beta_{3,1}\beta_{2,3}\beta_{3,2}$ (h) $\beta_{1,1}\beta_{3,1}\beta_{3,3}$ (i) $\beta_{2,1}\beta_{3,2}\beta_{3,3}$ (j) $\beta_{3,1}\beta_{3,3}\beta_{3,3}$ 

Note: Panel (2a) represents the total effect of η_X on η_Y at a time interval of three is a function of the lagged parameters for a time interval of one. Panels (2b) to (2j) represent the decomposition of the total effect (dashed path in Panel 2a) into direct (dashed paths) and indirect (solid paths) effects.

from marital conflict to behavioral competence without passing through maternal knowledge which is the direct effect, that is, -0.091 as well as the effects that pass from marital conflict to behavioral competence through maternal knowledge which is the indirect effect, that is, -0.009 . From this decomposition, we can see based on the absolute values of the direct and indirect effects, that much of the effect of maternal conflict on behavioral competence is direct but some effects pass through maternal knowledge.

Standardized Total, Direct, and Indirect Effects

Effect sizes in mediation models have become essential tools for quantifying the magnitude and practical significance of indirect effects, moving beyond the limitations of significance testing. Mediation models decompose the total effect of an independent variable on a dependent variable into direct and indirect effects, where the latter is transmitted through a mediator. Effect sizes offer a way to measure these relationships, enabling researchers to better understand and communicate the strength of mediated effects (Kelley & Preacher, 2012; Lachowicz et al., 2018; MacKinnon & Dwyer, 1993).

Several types of effect sizes have been proposed for mediation models, each with distinct advantages and limitations. Unstandardized indirect effects, calculated from raw regression coefficients, are straightforward to compute and interpret within a specific study, but are limited by their dependence on variable scales, which hinders cross-study comparisons (M. W.-L. Cheung, 2009a; Fairchild et al., 2009). Proportion mediated measures, which indicate the percentage of the total effect that is mediated, are particularly useful in models with multiple mediators. However, these measures can become unstable in small samples or create interpretational difficulties in some cases, such as when direct and indirect effects have opposite signs (M. W.-L. Cheung, 2009a; Fairchild et al., 2009). Similarly, R^2 -based measures provide insight into the variance explained by the mediator and the independent variable, but are sensitive to model specifications and the overall fit of the model (Fairchild et al., 2009; Preacher & Kelley, 2011).

Taking inspiration from the DT standardization proposed by Schuurman et al. (2016), we derive the standardized total, direct, and indirect effect for the CT mediation model as follows

$$\text{Total}_{i,j}^* = \text{Total}_{i,j} \left(\frac{\sigma_{x_j}}{\sigma_{y_i}} \right), \quad (13)$$

$$\text{Direct}_{i,j}^* = \text{Direct}_{i,j} \left(\frac{\sigma_{x_j}}{\sigma_{y_i}} \right), \quad \text{and} \quad (14)$$

$$\text{Indirect}_{i,j}^* = \text{Total}_{i,j}^* - \text{Direct}_{i,j}^* \quad (15)$$

where σ_{x_j} and σ_{y_i} are the steady-state model-implied standard deviations of the state independent and dependent variables, respectively, that is, the square root of the corresponding diagonal elements of the steady-state model-implied co-

variance matrix $\text{Cov}(\boldsymbol{\eta})$ which is given by solving for \mathbf{X} in the following equation:

$$\Phi \mathbf{X} + \mathbf{X}\Phi' + \Sigma = \mathbf{0}, \quad (16)$$

where $(\cdot)'$ represents the matrix transpose. This equation is known as the Lyapunov (1992) equation which provides a powerful tool to determine the steady-state covariance matrix in the context of linear SDEs. By solving the Lyapunov equation, we can characterize how the system's state will fluctuate around its equilibrium μ under the influence of noise¹⁰. Given the process noise covariance matrix

$$\Sigma = \begin{pmatrix} 0.095 & 0.001 & 0.007 \\ 0.001 & 0.103 & 0.001 \\ 0.007 & 0.001 & 0.099 \end{pmatrix}, \quad (17)$$

and the drift matrix Φ given earlier, the steady-state model-implied covariance matrix is given by

$$\text{Cov}(\boldsymbol{\eta}) = \begin{pmatrix} 0.701 & -0.195 & -0.142 \\ -0.195 & 0.841 & 0.403 \\ -0.142 & 0.403 & 0.930 \end{pmatrix}. \quad (18)$$

Given the values above, the standardized total effect, representing the sum of all standardized effects from marital conflict to behavioral competence, is quantified as -0.087 . This value reflects the cumulative impact observed by the third year (with $\Delta t = 3$ and time t scaled in years). Within this total, the standardized direct effect, -0.079 , accounts for the portion of the effect that bypasses maternal knowledge. Conversely, the standardized indirect effect, -0.008 , captures the pathways where maternal knowledge mediates the relationship¹¹. From this breakdown, it is evident that while the standardized direct effect constitutes the majority of the influence from marital conflict to behavioral competence, a smaller but notable portion operates through maternal knowledge.

Despite differences from the standard mediation framework, the CT mediation shares a key limitation: potential deviations from normality. In the standard framework, when both α and β are zero in the population, the distribution of the product term exhibits a kurtosis of 6, regardless of sample size—even asymptotically (Craig, 1936; MacKinnon, 2008). This heavy-tailed behavior is also evident in CT mediation models. To illustrate, we simulated a CT mediation model in which the indirect effect was set to zero. As shown in

¹⁰The solution to this equation can be numerically calculated using the `syl` function from the `Armadillo` C++ linear algebra library made available to R using the `RcppArmadillo` package (Eddelbuettel & Sanderson, 2014).

¹¹The `MedStd` function in the `cTMed` package calculates the standardized direct, indirect, and total effects for a specific time interval or a range of time intervals. A plot method for the `MedStd` function can be used to visualize the effects as a function of time intervals.

Figure 3, the sampling distributions of both the unstandardized indirect effect and the standardized effects exhibit substantial deviations from normality, particularly in the tails. Therefore, as in standard mediation, methods for quantifying uncertainty in CT mediation must be able to accommodate these non-normal distributions.

Illustration 2: Quantifying Uncertainty in Direct, Indirect, and Total Mediation Effects in CT-VAR Models

In the previous section, we showed how to calculate point estimates for the direct, indirect, and total effects for a given time interval from the CT-VAR drift matrix Φ . In particular, we know the direct, indirect, and total effects of marital conflict (η_X) on behavioral competence (η_Y) through maternal knowledge (η_M), for a time interval of 3, are -0.091 , -0.009 , and -0.100 , respectively. We have also provided an approach to standardizing the direct, indirect, and total effects as effect size measures given by -0.079 , -0.008 , and -0.087 , respectively. However, we still do not have a way to assess these effects from an inferential point of view.

In this manuscript, we focus on statistical inference for effects evaluated at specific time intervals (Δt). We do not test the CT model parameters directly (e.g., elements of the drift matrix), but instead focus on functions of these parameters that represent how effects accumulate or decay over time. Because these derived effects depend on the time interval, statistical significance is inherently time-dependent: at very short or very long Δt , the effects may approach zero—even if the underlying CT effect is nonzero (Driver, 2025). This behavior reflects how time allows effects to manifest (or dissipate). All confidence intervals presented correspond to effects evaluated at specific values of Δt .

In this section, we extend the previous section by providing methods for quantifying the uncertainty in the direct, indirect, and total effects we have defined. In the following sections, we introduce the delta, the MC, and the PB methods. We chose the delta and MC approaches because they are relatively easy to compute even for complex models and provide a contrast in inferential results when the assumption of symmetry is imposed on the sampling distribution of the indirect effect (as in the delta method) as compared to when asymmetry is accounted for (as in the MC method). For all analyses involving the MC method—including the empirical illustration and Monte Carlo simulation studies—we used 20,000 replications to ensure stable estimates. The PB method, which was used in the empirical illustration and the small-scale simulation study as a benchmark, was conducted using 1,000 replications due to its higher computational cost. We included the PB method as a comparison approach that takes into consideration the dependence in time series data. We discuss these methods in general and then apply them to the context of generating SEs and CIs as a way to quantify the uncertainty for the direct, indirect, and total effects.

The Delta Method

The delta method is a powerful statistical technique used to derive the asymptotic distribution of a random variable. It becomes particularly handy when dealing with functions of other random variables that are asymptotically Gaussian, that is, their distributions converge to a Gaussian distribution as sample size increases. As previously mentioned, it has been applied in psychological research to generate SEs and CIs for the indirect effect (Aroian, 1947; Goodman, 1960; Sobel, 1986, 1987). It has also been applied to other functions of parameters such as the standardized regression coefficients (Jones & Waller, 2015; Pesigan et al., 2023; K.-H. Yuan & Chan, 2011). Currently, popular SEM software packages such as Mplus (Muthén & Muthén, 2017) and lavaan (Rosseel, 2012) use the delta method as the standard approach to generating SEs and CIs for defined functions of SEM parameters. In R, the `Delta` and `DeltaGeneric` functions in the `betaDelta` (Pesigan et al., 2023) package can generate SEs and CIs for functions of parameters of fitted models as long as the estimates and the sampling variance-covariance matrix are provided.

We define the delta method following the multivariate central limit theorem. Let θ be $\text{vec}(\Phi)$, that is, the elements of the drift matrix Φ in vector form sorted column-wise. Let $\hat{\theta}$ be $\text{vec}(\hat{\Phi})$, that is, the estimates of the elements of the drift matrix. In the case of the standardized total, direct, and indirect effects θ and $\hat{\theta}$ consists of the vector of the elements of the drift matrix and the unique elements of the process noise covariance matrix. The uncertainty around any functions of the estimates for $\hat{\theta}$ is assumed to be characterized as:

$$\sqrt{n}(\mathbf{g}(\hat{\theta}) - \mathbf{g}(\theta)) \xrightarrow{D} \mathcal{N}(0, \mathbf{J}\boldsymbol{\Gamma}\mathbf{J}') \quad (19)$$

where \mathbf{g} is a vector of functions, \mathbf{J} is the matrix of first-order derivatives of the functions in \mathbf{g} with respect to the elements of θ and $\boldsymbol{\Gamma}$ is the asymptotic variance-covariance matrix of $\hat{\theta}$. From the former, we can derive the distribution of $\mathbf{g}(\hat{\theta})$ for a sample size n as follows

$$\mathbf{g}(\hat{\theta}) \approx \mathcal{N}(\mathbf{g}(\theta), n^{-1}\mathbf{J}\boldsymbol{\Gamma}\mathbf{J}'). \quad (20)$$

The uncertainty associated with the estimator $\mathbf{g}(\hat{\theta})$ is, therefore, given by $n^{-1}\mathbf{J}\boldsymbol{\Gamma}\mathbf{J}'$. When $\boldsymbol{\Gamma}$ is unknown, by substitution, we can use the estimated sampling variance-covariance matrix of $\hat{\theta}$, that is, $\hat{\boldsymbol{\Gamma}}(\hat{\theta})$ for $n^{-1}\boldsymbol{\Gamma}$. Therefore, the sampling variance-covariance matrix of $\mathbf{g}(\hat{\theta})$ is given by

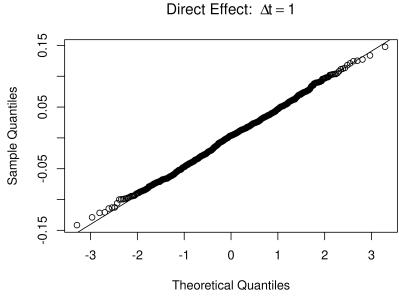
$$\mathbf{g}(\hat{\theta}) \approx \mathcal{N}(\mathbf{g}(\theta), \mathbf{J}\hat{\boldsymbol{\Gamma}}(\hat{\theta})\mathbf{J}'). \quad (21)$$

From this general approach, we can generate the sampling variance-covariance matrix of any function of the drift matrix Φ . The SEs, that is, the square root of the diagonal elements

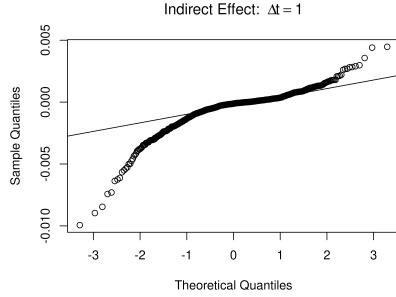
Figure 3

Normal QQ Plots and Effects Used in the Extended Monte Carlo Simulation Study

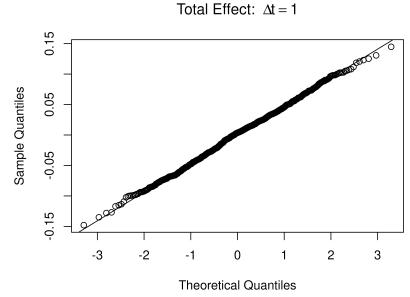
(a) Direct



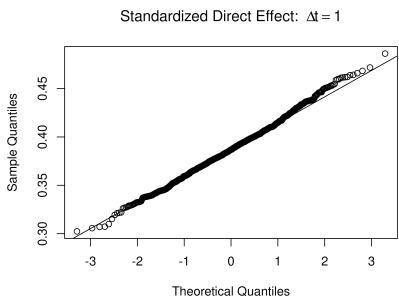
(b) Indirect



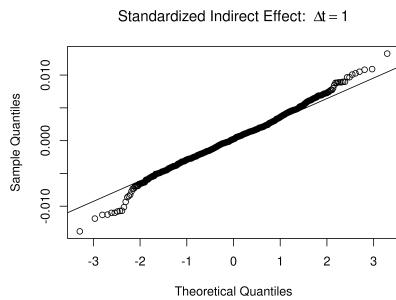
(c) Total



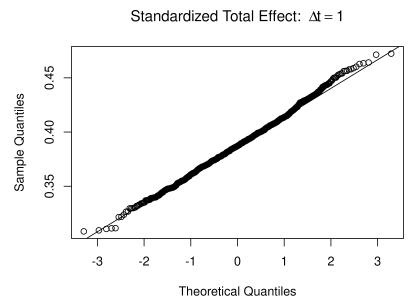
(d) Std. Direct



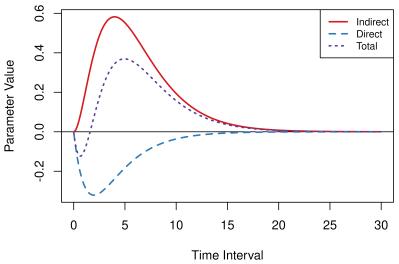
(e) Std. Indirect



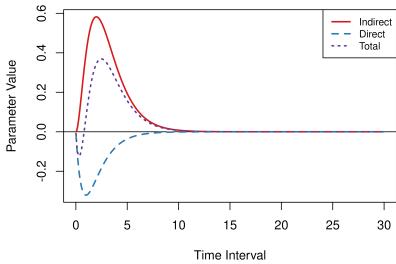
(f) Std. Total



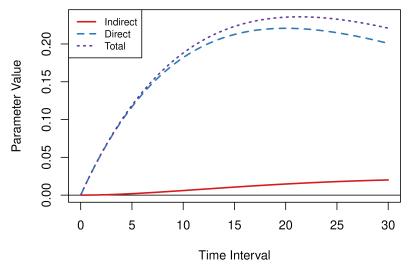
(g) Moderate Coupling (Φ_1)



(h) Strong Coupling (Φ_2)



(i) Weak Coupling (Φ_3)



Note: Deviations from the line in the QQ plots reflect deviation from normality. The scatter plots show the effects under different coupling strengths.

of the sampling variance-covariance matrix are used to generate symmetric CIs around the point estimates of any function of the drift matrix Φ . In the current application, let the function \mathbf{g} be a vector of the direct, indirect, and total effects of a specific independent variable on a dependent variable through one or more mediator variables. The `DeltaMed` and the `DeltaMedStd` functions in the `cTMed` package applies the delta method to generate SEs, test statistics, p -values, and CIs for unstandardized and standardized effects, respectively, for a specific time interval or a range of time intervals.

The Monte Carlo (MC) Method

MC was introduced in the mediation literature by MacKinnon et al. (2004) and formalized by Preacher and Selig (2012) as a resampling technique for generating CIs for the indirect effect. It has been shown to be as good as NB (Pesigan & Cheung, 2023; Tofghi & Kelley, 2019, 2020; Tofghi & MacKinnon, 2015). As a resampling technique, it is similar to the bootstrap such that the sampling distribution of the function of parameters does not need to be normal. Unlike the bootstrap, however, it is more computationally efficient as it does not involve multiple model fitting. The R package

`semcci` provides a general approach to applying MC for generating CIs for a function of any SEM parameters. The `MCFunc` and `MCGeneric` functions in the `semcci` (Pesigan & Cheung, 2023) package can generate SEs and CIs for functions of parameters of fitted models as long as the estimates and the sampling variance-covariance matrix are provided.

Based on the asymptotic properties of maximum likelihood estimators, we can assume that they are normally distributed around the population parameters

$$\hat{\theta} \sim N(\theta, \mathbb{V}(\theta)). \quad (22)$$

Using this distributional assumption, a sampling distribution of $\hat{\theta}$ which we refer to as $\hat{\theta}^*$ can be generated by replacing the population parameters with sample estimates, that is,

$$\hat{\theta}^* \sim N(\hat{\theta}, \hat{\mathbb{V}}(\hat{\theta})). \quad (23)$$

A sampling distribution of $\mathbf{g}(\hat{\theta})$, which we refer to as $\mathbf{g}(\hat{\theta}^*)$, can be generated by using the simulated estimates (e.g., $R = 20,000$) to calculate \mathbf{g} . The standard deviations of the simulated estimates are the SEs. In resampling approaches, it is customary to use the percentiles of the generated distribution as CIs instead of the SEs. This approach allows the confidence limits to be asymmetric around the point estimates, unlike the delta method which is always symmetric. Percentiles corresponding to $100(1 - \alpha)\%$ are the CIs where α is the significance level (e.g., 0.05).

From this general approach, we can generate the empirical sampling distribution of any function of the drift matrix Φ . There are constraints, however, on the randomly generated values of Φ . First, the diagonal values should be from zero to negative infinity. Second, the real part of all eigenvalues is less than zero. These two can be implemented using a replacement sampling approach, that is, discarding randomly generated values where the two constraints do not hold. In the case of the standardized total, direct, and indirect effects, we also generate random values for the positive definite process noise covariance matrix Σ . Again, let the function \mathbf{g} be a vector of the direct, indirect, and total effects of a specific independent variable on a dependent variable through one or more mediator variables. While MC follows the maximum likelihood asymptotic assumption by generating CT estimators, θ , from a multivariate normal distribution (subject to necessary constraints), it does not assume that the non-linear transformation of these estimators remains normally distributed. Instead, MC allows for the sampling distribution of the transformed parameters to be empirically derived, capturing potential skewness and non-normality that arises from the CT-to-DT transformation. This is a key advantage of MC over the delta method, as it remains robust even when the transformed distributions exhibit non-normal characteristics. The `MCMed` and `MCMedStd` functions in the `cTMed` package applies MC to generate the SEs and CIs for unstandardized

and standardized effects for a specific time interval or a range of time intervals.

Parametric Bootstrap (PB) Method

We performed targeted comparisons of the delta and MC methods to a PB approach, which has been shown to perform well for intensive longitudinal data under conditions with correctly specified models (Zhang, 2018; Zhang & Browne, 2010). The PB approach considered meets two critical conditions: they preserve the data dependence characteristic of multivariate time series, and they are generated in accordance with the model of interest, as highlighted by Lahiri (2003). The proposed PB method begins by estimating the parameters of the CT-VAR model from the observed data using tools such as `dynr` or `OpenMx`. These estimated parameters are then used to generate new datasets that simulate the temporal structure of the original data, a process facilitated by the `PBSSMOUFixed` function from the `bootStateSpace`¹² package using Equations 1 and 3 where the CT-VAR model is converted to the DT solution given by Equation 5 and a small Δt (e.g., 0.10) is used to generate smooth state trajectories. Once the simulated datasets are created, the CT-VAR model is re-estimated using the `dynr` package for each dataset to produce new parameter estimates. From these re-estimated parameters, the total, direct, and indirect effects are derived using the `BootMed` function for the unstandardized effects and the `BootMedStd` function for standardized effects, from the `cTMed` package. This sequence of data generation, model re-estimation, and effect derivation is repeated B times (e.g., $B = 1,000$) to create an empirical sampling distribution of the effects.

The empirical sampling distribution obtained through this process enables the computation of key inferential statistics. Standard errors are calculated as the standard deviation of the bootstrap estimates across repetitions. Confidence intervals are derived using the percentile (PC) method, which identifies the appropriate quantiles corresponding to the desired level of confidence. Additionally, bias-corrected (BC) estimates are obtained by adjusting for any systematic differences between the original and bootstrap estimates.

CIs for the Direct, Indirect, and Total Effects

Using the delta, MC, and PB methods as implemented by the `cTMed` package, the SEs and CIs for the unstandardized and standardized direct, indirect, and total effects of marital conflict (η_X) on behavioral competence (η_Y) through maternal knowledge (η_M), for time intervals Δt of 1, 2, and 3, are presented in Table 2. Based on these results, we highlight several key observations. First, the magnitude of the direct, indirect, and total effects increased with longer time

¹²See <https://CRAN.R-project.org/package=bootStateSpace>. The specific version used is 1.0.2.

intervals, suggesting stronger effects at larger temporal lags. However, at sufficiently large time intervals, the effects eventually taper off toward zero. All three methods produced confidence intervals (CIs) for the total and direct effects that did not overlap with zero across $\Delta t = 1, 2$, and 3 , indicating statistical significance regardless of the method used. In contrast, the indirect effect CIs overlapped with zero for all three methods and all time intervals, indicating non-significance. While the delta method produced symmetric CIs, both the MC and PB methods accounted for the asymmetry typically observed in mediation effects. This makes MC and PB more appropriate when the sampling distribution of the effect is non-normal, which is often the case in indirect effects. Importantly, the MC and PB methods yielded comparable results in this example. However, the MC method has a major computational advantage: it completes in a matter of minutes, whereas the PB method can take several hours to run. Thus, while both are appropriate in contexts where asymmetry is present, the MC method provides a practical, computationally efficient alternative to the PB method, especially in large-scale or time-sensitive applications.

Illustration 3: Using Regions of Significance in Direct, Indirect, and Total Effects to Inform Designs of Longitudinal Mediation Studies

The Johnson-Neyman technique (Bauer & Curran, 2005; Johnson & Fay, 1950; Johnson & Neyman, 1936; Preacher et al., 2006) is a statistical method utilized for analyzing interactions in regression models. It addresses the limitations of the simple slopes method by offering a more systematic approach to determining the *regions of significance* for interaction effects. Rather than selecting arbitrary conditional values of the moderator variable, this technique computes the range of values of the moderator that result in a significant difference in the relationship between the predictor and the outcome variable. By computing the conditional values needed to yield test statistics (e.g., t -statistic) that exceed a pre-determined critical threshold (typically to achieve a p -value of .05), the Johnson-Neyman technique delineates the regions of significance for the moderator variable, which signify the range of moderator values within which the relationship between the predictor and the outcome variable is statistically significant.

We aim to extend this concept from the regression framework to the CT mediation. Specifically, we seek to identify regions of significance for the direct, indirect, and total effects—that is, the time interval values for which these effects are statistically significant. As noted earlier, the parameters in the DT solution of CT-VAR are defined for any time interval. As such, we can generate a spectrum of time intervals Δt , and compute the corresponding direct, indirect, and total effects using the CT-VAR parameter estimates. Using the delta, MC, and PB methods, CIs can be generated to

depict the regions of significance for the direct, indirect, and total effects for these time interval values, as analogous to the Johnson-Neyman technique. Whereas previous research has focused on point estimates and finding the maximal effects (e.g., Deboeck & Preacher, 2015; Hecht & Zitzmann, 2021; Wang & Zhang, 2020), the methods we have developed in previous sections adds the ability to visualize regions where the direct, indirect, and total effects are significantly different from zero.

Figure 4 present the regions of significance for the unstandardized direct, indirect, and total effects, respectively, for the range of time intervals > 0 to 10 . The direct and total effects are statistically significant across time intervals for PB PC and PB BC. The indirect effect, however, overlaps with zero, indicating a lack of statistical significance.

A sole reliance on significance testing can lead to misinterpretation, as statistical significance is heavily influenced by sample size, model specification, and noise in the data. Instead of interpreting results purely in terms of whether an effect is statistically significant, researchers should adopt a dual perspective that considers both effect size magnitude and uncertainty quantification through CIs. The standardized direct, indirect, and total effects provide an intuitive measure of the strength of mediation effects independent of measurement units, offering insight into the practical significance of findings. Even if an effect is statistically significant, its substantive importance should still be assessed based on its effect size. Conversely, non-significant effects do not necessarily imply an absence of mediation; rather, they may indicate greater uncertainty, which can be examined through the width of the confidence interval. A narrow CI suggests more precision, while a wide CI indicates greater variability, helping researchers determine whether an effect is truly small or if the study lacks the statistical power to detect it reliably. By incorporating both effect size interpretation and CIs, this approach ensures that mediation effects are evaluated not only in terms of statistical significance but also in terms of practical relevance. This perspective prevents over-reliance on binary hypothesis testing and promotes a more informed, meaningful interpretation of CT mediation effects.

Monte Carlo Simulation Study

To evaluate the performance of our proposed methods for quantifying uncertainty in CT mediation, we conducted two simulation studies. The first is a small-scale simulation study, designed to benchmark the delta and MC methods against the PB using parameters estimated from an empirical dataset. This allowed us to assess performance under realistic, empirically-motivated conditions.

We then conducted an extended simulation study to systematically examine the robustness of the delta and MC methods across a broader range of conditions, including varying sample sizes, time intervals, and effect sizes. This

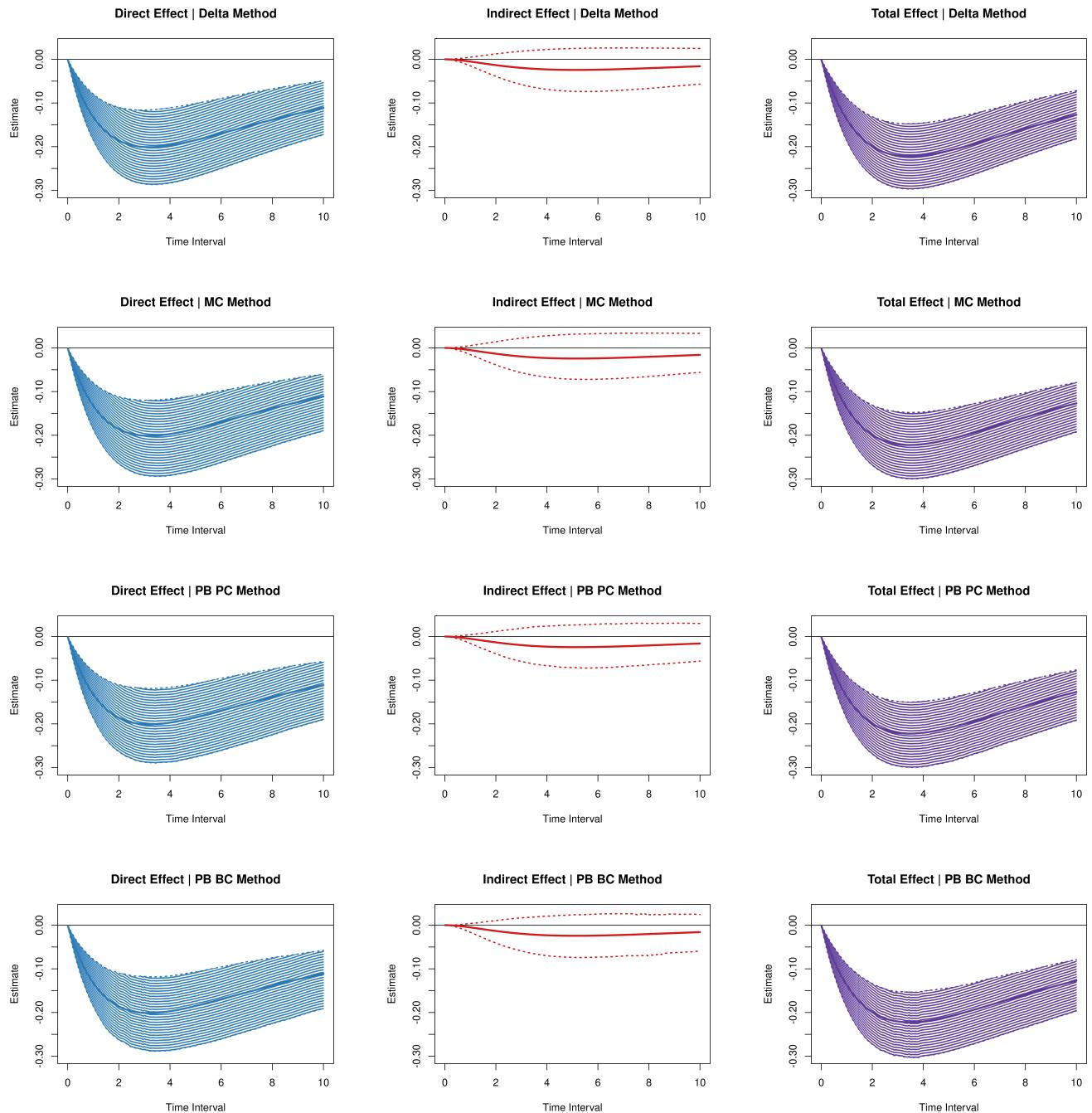
Table 2
Point Estimates, SEs, and 95% CIs of the Direct, Indirect, and Total Effects for a Single Data Set Simulated from Grundy et al. (2007)

Unstandardized																
Δt	Effect	Estimate	Delta			MC			PB PC							
			SE	2.5%	97.5%	SE	2.5%	97.5%	SE	2.5%	97.5%	Sym*				
1	Direct	-0.060	0.012	-0.084	-0.036	0.000	0.012	-0.084	0.000	0.013	-0.086	-0.036	0.001			
	Indirect	-0.002	0.002	-0.007	-0.004	0.000	0.002	-0.007	0.002	0.002	-0.007	0.002	0.001			
	Total	-0.063	0.011	-0.085	-0.040	0.000	0.011	-0.085	-0.040	0.001	-0.086	-0.040	0.001			
2	Direct	-0.085	0.017	-0.119	-0.050	0.000	0.018	-0.120	-0.051	0.001	-0.121	-0.050	0.002			
	Indirect	-0.006	0.006	-0.018	-0.006	0.000	0.006	-0.017	0.006	0.001	-0.018	0.005	0.001			
	Total	-0.091	0.016	-0.121	-0.060	0.000	0.016	-0.121	-0.059	0.001	-0.121	-0.061	0.002			
3	Direct	-0.091	0.019	-0.129	-0.054	0.000	0.019	-0.131	-0.055	0.003	-0.131	-0.053	0.003			
	Indirect	-0.009	0.009	-0.026	0.008	0.000	0.009	-0.026	0.010	0.002	-0.026	0.008	0.001			
	Total	-0.100	0.017	-0.133	-0.067	0.000	0.017	-0.133	-0.067	0.001	-0.134	-0.067	0.002			
Standardized																
Δt	Effect	Estimate	Delta			MC			PB PC							
			SE	2.5%	97.5%	SE	2.5%	97.5%	SE	2.5%	97.5%	Sym*				
1	Direct	-0.132	0.027	-0.185	-0.080	0.000	0.027	-0.186	-0.080	0.002	-0.187	-0.079	-0.079	0.000		
	Indirect	-0.005	0.005	-0.015	-0.005	0.000	0.005	-0.015	0.005	0.000	-0.016	-0.016	0.004	0.002		
	Total	-0.137	0.025	-0.187	-0.088	0.000	0.025	-0.188	-0.088	0.001	-0.189	-0.089	0.003	0.003		
2	Direct	-0.186	0.039	-0.262	-0.110	0.000	0.039	-0.266	-0.112	0.005	-0.263	-0.110	0.002	-0.110	0.000	
	Indirect	-0.013	0.013	-0.039	0.012	0.000	0.013	-0.039	0.014	0.002	-0.039	0.013	0.011	0.004		
	Total	-0.199	0.035	-0.267	-0.131	0.000	0.035	-0.269	-0.131	0.001	-0.269	-0.132	0.002	-0.134	0.008	
3	Direct	-0.201	0.043	-0.286	-0.116	0.000	0.044	-0.292	-0.120	0.010	-0.288	-0.118	0.004	-0.118	0.001	
	Indirect	-0.020	0.019	-0.058	0.019	0.000	0.020	-0.057	0.022	0.005	0.019	-0.057	0.019	-0.059	0.017	0.003
	Total	-0.221	0.038	-0.294	-0.147	0.000	0.038	-0.297	-0.146	0.002	0.039	-0.296	0.004	-0.299	-0.152	0.010

Note. *The symmetry of the CIs is assessed using the absolute value of $2.5\% - \text{estimate} + 97.5\% - \text{estimate}$. Values greater than zero indicate asymmetry.

Figure 4

Regions of Significance for the Standardized Direct, Indirect, and Total Effects (Conflict → Knowledge → Competence)



Note: The shaded areas represent regions of significance, that is, instances where the 95% CIs did not contain zero for a given time interval.

broader evaluation provides a general understanding of each method's behavior in applied settings.

Simulation Study 1: Empirical Benchmark Against the Parametric Bootstrap

Since the methods proposed in the empirical illustration yielded similar results, we conducted a small-scale Monte Carlo simulation, generating 1,000 datasets based on the parameters reported in Grundy et al. (2007). The goal was to compare the performance of the delta and MC methods to the PB method using data simulated to reflect the empirical illustration. See Figure 5 and 6 for plots of the results.

We evaluated coverage probabilities—the proportion of simulation replications in which the CIs contained the population parameter—to assess the performance of each method. Using a 95% CI, we deemed coverage within the range of 0.925 to 0.975 as robust (Bradley, 1978). Both the MC and PB methods achieved coverage close to the nominal level, while the delta method showed a noticeable deviation. For both direct and total effects, the delta method's coverage probability declined as the time interval increased. Interestingly, the opposite trend was observed for the indirect effect, where coverage increased as the time interval grew. These results suggest that MC and PB produce reliable CIs, whereas the delta method tends to yield biased CIs as the time interval increases.

We also examined statistical power, considering values of 0.80 or higher as robust (Cohen, 1988). For direct and total effects, both the MC and PB methods achieved power above this threshold, while the delta method's power declined as the time interval increased. All methods showed low power for detecting the indirect effect, likely due to its small effect size, which typically requires larger samples. Among the approaches for the indirect effect, PB BC was the most powerful, followed by MC and PB PC, with the delta method performing the worst overall.

Simulation Study 2: Extended Monte Carlo Simulation of Method Performance

The small-scale simulation in the preceding section provided a useful benchmark for comparing the delta and MC methods against the PB approach. The results showed that MC performs comparably to PB, suggesting that MC may be a viable and more computationally efficient alternative in practice. Building on this preliminary evidence, we conducted an extended simulation study—focusing exclusively on the delta and MC methods—to evaluate their performance under a broader range of conditions. We did not include the PB method in this extended study due to its prohibitive computational demands: while the delta method completes almost instantaneously and the MC method typically takes only a few minutes, a single Monte Carlo simulation replica-

tion using PB can take up to 36 hours on a single core of a Dell Latitude 5340 laptop.

We compared the delta and MC approaches across varying sample sizes (ranging from 50 to 500), different time intervals (Δt values from 1 to 30), and various magnitudes of direct, indirect, and total effects resulting from those time intervals. The model parameters were based on the “benchmark” example from Deboeck and Preacher (2015), which features a suppression effect or inconsistent mediation (MacKinnon et al., 2000; Shrout & Bolger, 2002), where the direct and indirect effects carry opposite signs. Our primary focus was to assess how the delta and MC methods perform in quantifying uncertainty within this mediation context, using metrics such as coverage probability, statistical power, and Type I error rates for the CIs.

Parameters

Deboeck and Preacher (2015) generated data from a DT-VAR model using a fixed time interval of one unit with the following parameters

$$\boldsymbol{\alpha}_{\Delta t(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\boldsymbol{\beta}_{\Delta t(1)} = \begin{pmatrix} 0.70 & 0.00 & 0.00 \\ 0.50 & 0.60 & 0.00 \\ -0.10 & 0.40 & 0.50 \end{pmatrix}, \text{ and}$$

$$\boldsymbol{\Psi}_{\Delta t(1)} = \begin{pmatrix} 0.10 & 0.00 & 0.00 \\ 0.00 & 0.10 & 0.00 \\ 0.00 & 0.00 & 0.10 \end{pmatrix}$$

and initial values of η_X , η_M , and η_Y sampled from a standard multivariate normal distribution. Using the DT-VAR parameters above, we derived the CT-VAR parameters following the approach proposed by Chow et al. (2023), arriving at the following

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\boldsymbol{\Phi}_1 = \begin{pmatrix} -0.36 & 0.00 & 0.00 \\ 0.77 & -0.51 & 0.00 \\ -0.45 & 0.73 & -0.69 \end{pmatrix}, \text{ and}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.24 & 0.02 & -0.05 \\ 0.02 & 0.07 & 0.02 \\ -0.05 & 0.02 & 0.08 \end{pmatrix}.$$

We used the same mean vector and covariance matrix to generate the initial values of η_X , η_M , and η_Y . In addition to Deboeck and Preacher (2015), we included a measurement model for the observed variables with the following param-

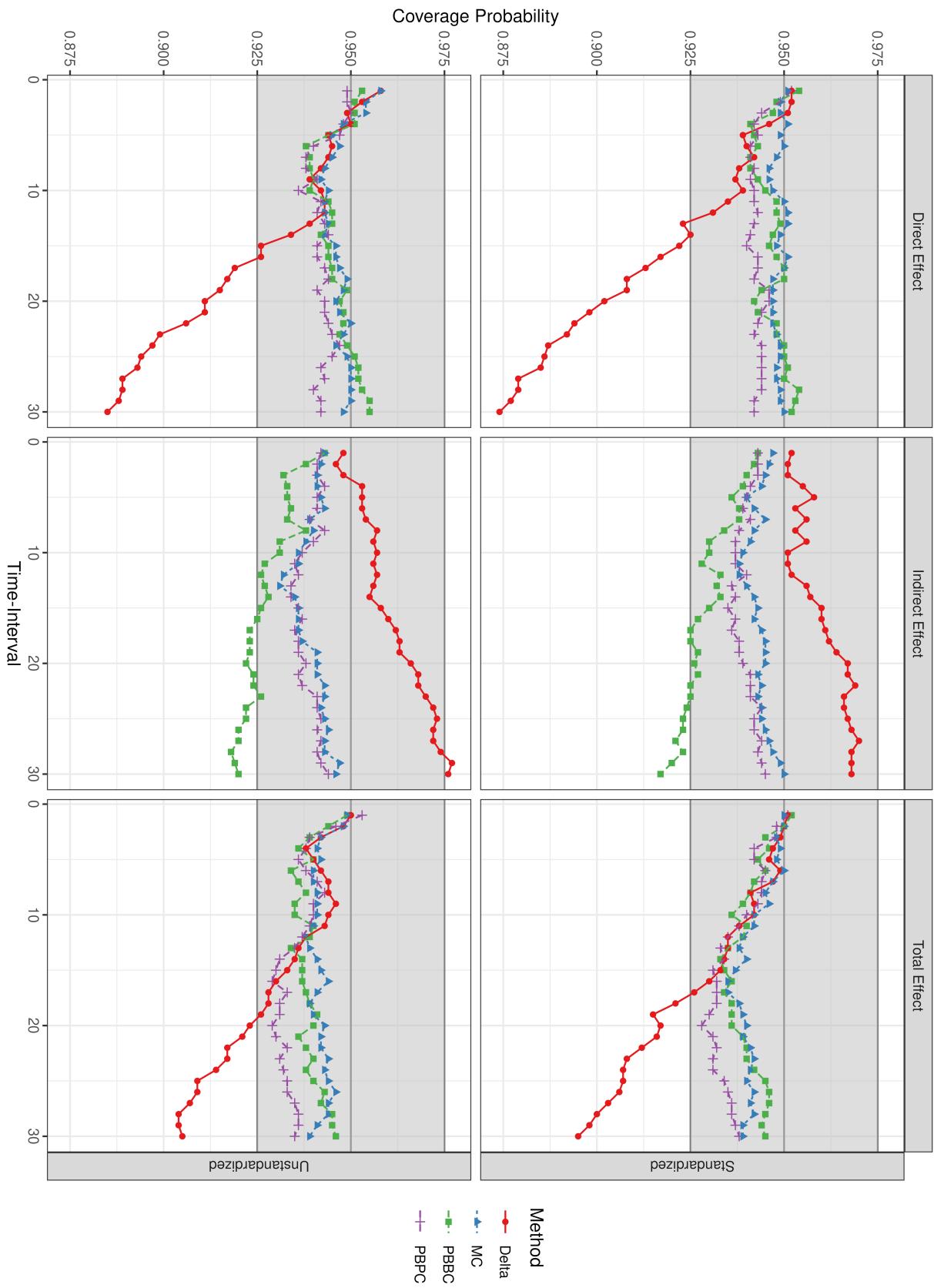
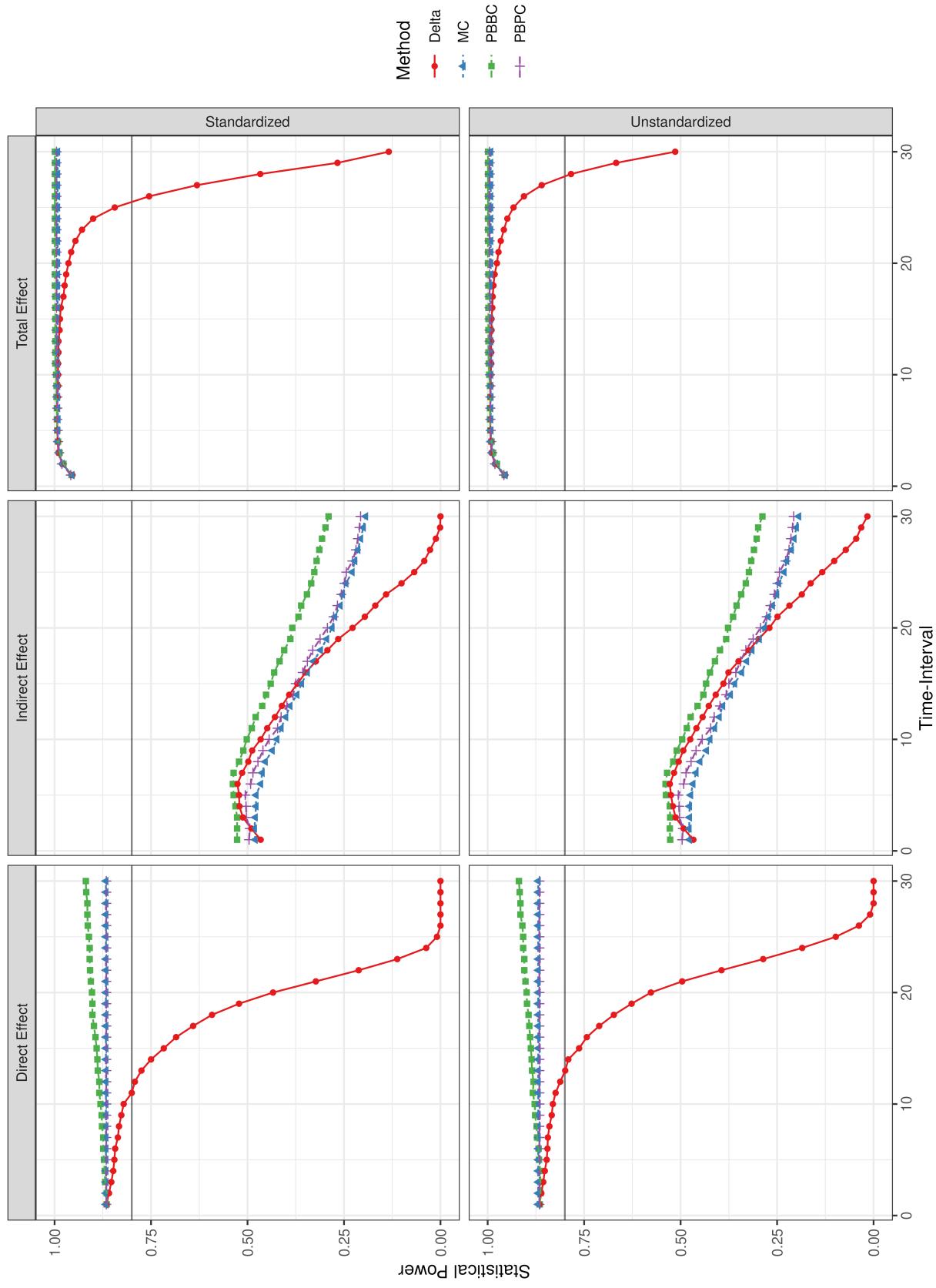
Figure 5*Coverage Probabilities*

Figure 6
Statistical Power



ters

$$\nu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and}$$

$$\Theta = \begin{pmatrix} 0.20 & 0.00 & 0.00 \\ 0.00 & 0.20 & 0.00 \\ 0.00 & 0.00 & 0.20 \end{pmatrix}.$$

To capture diverse dynamics, we included two additional drift matrices:

$$\Phi_2 = \begin{pmatrix} -0.71 & 0.00 & 0.00 \\ 1.54 & -1.02 & 0.00 \\ -0.90 & 1.46 & -1.39 \end{pmatrix}, \text{ and}$$

$$\Phi_3 = \begin{pmatrix} -0.05 & 0.00 & 0.00 \\ -0.02 & -0.05 & 0.00 \\ 0.03 & -0.01 & -0.05 \end{pmatrix}.$$

These matrices represent three dynamic regimes: (1) Φ_1 : moderate coupling (intermediate decay, meaningful interactions), (2) Φ_2 : strong coupling (rapid decay, intensified interactions), and (3) Φ_3 : weak coupling (slow decay, minimal interactions). This variability enables a comprehensive examination of system dynamics, from tightly coupled fast-evolving processes to near-independent slow evolution, ensuring robust and generalizable insights into system behavior. The process noise covariance matrices were chosen such that the model-implied state covariance matrix for the three cases are similar. The direct, indirect, and total effects that correspond to the drift matrices Φ_1 , Φ_2 , and Φ_3 are plotted in Figure 3.

Data Generation

We simulated values for three variables— X , M , and Y for individual i at time t and measurement occasion m . Varying the sample size n (50 to 500), we explored the impact on CI methods, with a focus on the delta method, which requires a large sample size. Measurement occasions spanned daily measurements over 180 days, represented by the vector $m = \{0, \dots, 179\}$. The vector of time is given by $t = m \times \Delta t$, where $\Delta t = \frac{1}{30}$ (equivalent to a day in a month). Irregularly spaced data resulted from randomly selecting 100 observations per individual and flagging unselected observations as missing values. Latent variables η_X , η_M , and η_Y were generated using the DT integral solution of CT-VAR (Equation 5), and observed variables followed the measurement model (Equation 1). Data was simulated using the wrapper function `manCTMed::GenData` which utilized the `SimSSMOUFixed` function from the `simStateSpace` R package.

Model Fitting and CIs

The CT-VAR model parameters were estimated using the `manCTMed::FitDynd` wrapper function, which utilized the `dynd` package. Specifically, the Λ matrix was fixed as a diagonal matrix, the μ vector was set to a null vector, and the off-diagonal elements of the Θ matrix were fixed at zero while the diagonal elements were estimated. The remaining parameters were estimated as well.

From the fitted model, we extracted the estimated Φ matrix and Σ matrix and their corresponding sampling variance-covariance matrix. Using these, we generated CIs via delta and MC methods for time intervals ranging from 1 to 30. We considered two mediation models: one involving the path $\eta_X \rightarrow \eta_M \rightarrow \eta_Y$ and the other involving $\eta_Y \rightarrow \eta_M \rightarrow \eta_X$. In the first model, although effects approach zero at higher time intervals, we anticipate non-zero direct, indirect, and total effects within the range of Δt from 1 to 30. Conversely, in the second model, we expect all effects to be zero across all Δt values. The `manCTMed::Delta*` and `manCTMed::MC*` wrapper functions employed the `DeltaMed`, `DeltaMedStd`, `MCMed`, and `MCMedStd` functions from the `cTMed` package to generate these CIs.

Monte Carlo Simulation Study Results for the Expanded Comparisons

In addition to evaluating statistical power and coverage for the mediation model $\eta_X \rightarrow \eta_M \rightarrow \eta_Y$, we assessed Type I error rates using the reversed model $\eta_Y \rightarrow \eta_M \rightarrow \eta_X$, where all effects were set to zero across time intervals. Type I error reflects the proportion of replications in which the CI excluded zero despite the true parameter being zero—that is, the probability of falsely rejecting a true null hypothesis. Plots of the complete results are available in the supplementary materials. While there are some differences in the results for the moderate, strong, and weak coupling, the pattern of results when it comes to the comparison of the performance of the delta and MC methods are similar across the different drift matrices. We will focus our attention of the results for the standardized effects with the moderate coupling case.

Coverage Probabilities. For Δt values from 1 to 4, both direct and total effects met the robustness criteria in terms of coverage probabilities using the delta method. However, beyond Δt values greater than 4 but less than 28, the delta method encountered coverage issues, particularly affecting the direct effect. The CIs became inflated, resulting in coverage probabilities equal to one. In contrast, MC consistently maintained robust coverage for indirect and total effects across various Δt values, even exceeding 28. Notably, the direct effect's coverage dropped significantly at Δt exceeding 28, where the effect size was extremely small. Detecting effects at large time intervals may be practically challenging due to their diminished magnitude and reduced statistical power; however, such effects may still carry theoret-

ical importance depending on the research context. Overall, MC outperformed the delta method, except at Δt values greater than 28 for the direct effect. Additionally, sample size n influenced the delta method's performance: large sample sizes achieved robust CIs for smaller Δt values, but even with a sample size of 500, robustness was not achieved for larger Δt . In contrast, MC remained robust regardless of sample size, except for extremely large Δt where the effect size approached zero. In summary, MC was the better choice for CI coverage.

Statistical Power. As Δt increased, the absolute effect size approached zero, leading to an expected decrease in statistical power. Notably, the total effect exhibited robust power (≥ 0.8) for Δt values from 1 to 4. However, at $\Delta t = 8$, the delta method's power declined below 0.8, especially with smaller sample sizes. MC also experienced reduced power starting at $\Delta t = 12$. Overall, while both methods showed decreasing power with increasing Δt , the delta method declined earlier and faster than MC. Sample size n played a role: MC demonstrated improved power as n increased, surpassing the delta method. For indirect and total effects, MC consistently outperformed the delta method. In summary, MC tends to produce more powerful confidence intervals than the delta method, making it a preferable option in practice.

Type I Error Rate. In the delta method, Type I error rates remained robust for direct and total effects within Δt values from 1 to 6. However, beyond $\Delta t = 6$, the delta method's error rates deflated to zero. MC exhibited distinct behavior: while some cases deviated from the robustness criteria (especially at higher Δt values), rates remained close to the nominal level of 0.05. Notably, MC maintained robustness even with increasing Δt . Interestingly, the method behaved differently for indirect effects compared to direct and total effects. The delta method consistently yielded a deflated Type I error rate of zero, whereas MC's rates, though not entirely deflated, did not meet the robustness criteria. Prior research by M. W.-L. Cheung (2009a) and Pesigan and Cheung (2023) reported similar patterns. Specifically, when both coefficients α and β were true zeros in a simple mediation model, the resulting Type I error rate tended to be lower than the nominal α level. Considering all factors, MC emerged as the more robust choice for constructing CIs, particularly when addressing Type I errors.

Conclusions from the Monte Carlo Simulation Study

To evaluate the finite-sample performance of different CI estimation methods in CT mediation models, we conducted two simulation studies. The first, small-scale simulation benchmarked the delta and MC methods against PB, using data simulated to reflect an empirical illustration from Grundy et al. (2007). Results showed that the MC method closely matched the performance of PB in terms of both coverage probability and statistical power for direct and total ef-

fects. In contrast, the delta method's performance declined as the time interval increased, particularly in its ability to maintain accurate coverage.

Building on these findings, we conducted an extended Monte Carlo simulation study to more systematically compare the delta and MC methods across a broader range of conditions. Specifically, we varied sample sizes, time intervals, and effect sizes to assess the robustness of each method. The MC method consistently achieved robust coverage for indirect and total effects, even at large time intervals (e.g., exceeding 28 units). However, its coverage for the direct effect began to decline at intervals greater than 28. The delta method, in comparison, was less robust overall and more sensitive to increases in time interval length. These results reinforce the conclusion that the MC method is a reliable and preferred approach for constructing confidence intervals for direct, indirect, and total effects in CT mediation models.

Using the delta method to approximate the SEs and CIs for the direct, indirect, and total effects that involve matrix exponentials can be convenient but comes with several potential pitfalls. The primary concern is that the matrix exponential is highly nonlinear, and the delta method relies on a first-order Taylor expansion, which may not provide an accurate approximation, particularly when the elements of drift matrix Φ are large. Since the matrix exponential depends on the eigenvalues of Φ , even small changes in Φ can lead to significant variations in $\exp(\Phi)$, especially if Φ has complex or near-zero eigenvalues. This sensitivity means that the delta method's normal approximation may underestimate or mischaracterize the true variability in $\exp(\Phi)$, leading to CIs that do not adequately capture the uncertainty.

Another issue is that the variance propagation in the delta method is inherently asymmetric and ignores higher-order terms, which can lead to incorrectly centered CIs. Additionally, the estimated covariance matrix of $\exp(\Phi)$ is not necessarily positive semi-definite, which could result in negative variance estimates or invalid CIs, particularly when dealing with small sample sizes or high-dimensional matrices. The uncertainty in Φ can propagate in nonlinear ways that are not well captured by a simple variance approximation.

Given these limitations, alternative approaches may provide more robust uncertainty quantification. Bootstrapping can offer more accurate standard errors and CIs, particularly when sample sizes are small or Φ has high variability. We introduced a parametric approach, however, the computational intensity restricted us including it in the Monte Carlo simulation study. We proposed MC which is less computationally expensive as PB and showed that its performance is superior to the delta method.

Discussion

Taking inspiration from the mediation literature, we have extended the work of Deboeck and Preacher (2015) and Ryan

and Hamaker (2021) by proposing the delta, MC, and PB methods as ways to quantify the uncertainty in the point estimates of unstandardized and standardized direct, indirect, and total effects in CT mediation. We have shown how we build on and extend previous work through the three illustrations using a single empirical data set. Furthermore, we discussed the implications of our contributions to substantive research. We have shown through the simulation study that the MC CI is superior to the delta method in terms of coverage, power, and Type I error. We have also provided a free and open-source R package *cTMed* to implement the methods we have proposed together with miscellaneous functions to make the interpretation and presentation of the results easier. Of particular importance is the plotting capabilities of the package which facilitates visualizing how effects change over time and regions of statistical significance with respect to the time interval. The package will be a valuable tool for applied researchers in using the methods we proposed in their substantive area of research.

In our study, we made specific data analytic decisions to address key issues. First, we estimated the measurement model to account for measurement error. Using single indicators for each latent variable with fixed factor loadings of one and freely estimated residual variances, we avoided the common practice of ignoring the measurement model, which can impact parameter estimates in the dynamic structure (Oh et al., 2025; Schuurman & Hamaker, 2019; Schuurman et al., 2015; Staudenmayer & Buonaccorsi, 2005). Additionally, we handled irregularly spaced data by treating it as a missing data problem within the context of the CT framework. We employed an oversampling technique (Singer, 2012; Voelkle & Oud, 2012) where a small time interval Δt was chosen, and missing data were inserted between observed cases to aid in fitting CT models. While this improved convergence, it also extended the optimization process due to the insertion of missing values. Despite this solution, other numerical challenges persisted. In *dynr*, the SDEs are solved via numerical integration as opposed to using direct analytic solutions. Doing so allows for the possibility of fitting nonlinear SDEs, but this approach could also create more numerical problems in the estimation and Hessian computation process.

Limitations and Future Directions

We acknowledge some of the limitations of our work and propose future directions. In our Monte Carlo simulation study, we manipulated the sample size n , while keeping the number of measurement occasions fixed at 180. The length of the time series has the potential to impact the point estimates and SEs of the CT-VAR model. These estimates, in turn, affect the CIs for direct, indirect, and total effects. Future research could investigate the performance of delta and MC methods on data with varying measurement occasions.

The current study employed a frequentist approach for modeling CT-VAR using the *dynr* package. In the current approach, the parameters are fixed and do not vary across individuals. A multilevel mediation perspective where the repeated measures are nested within an individual (Kenny et al., 1998) is a logical extension of CT mediation. In this approach, coefficients for each individual have fixed and random components. Work on the Bayesian mediation model (e.g., Enders et al., 2013; Y. Yuan & MacKinnon, 2009) makes multilevel mediation easier compared to the frequentist approach. A Bayesian hierarchical or multilevel CT-VAR model is a more flexible alternative to the current approach (Driver & Voelkle, 2018; Hecht & Zitzmann, 2020; Oravecz et al., 2011, 2018). In the Bayesian CT-VAR framework, distributions of the direct, indirect, and total effects can be generated by calculating them from the posterior distribution of the drift matrix Φ . Credible intervals using the appropriate percentiles can be obtained from these distributions. To implement this approach, we can use the *PosteriorMed* function in the *cTMed* package. Alternatively, we can treat the posterior mean vector and covariance matrix of the posterior distribution of the drift matrix Φ as frequentist estimates which can then be used in the delta and MC methods. Further exploration and investigation of these ideas are warranted to better understand their implications and performance.

Our study considered a perspective where all variables are interconnected. While this works for small models, it becomes problematic for large, high-dimensional systems due to parameter explosion. To address this, we recommend using regularization techniques, specifically the lasso introduced by Tibshirani (1996, 2011). The lasso method shrinks parameters, introducing sparsity in the drift matrix Φ (Orzek & Voelkle, 2023). The implications of this regularization approach extend to calculating direct, indirect, and total effects in high-dimensional systems. Notably, parameters that are reduced to zero by lasso can be treated as fixed when applying the delta and MC methods. However, to validate the performance of these adjustments in practice, further investigation is needed.

In our current work, we adopted a path-tracing perspective, focusing on clarifying the temporal relationships between variables. However, we did not explore the network analysis perspective, such as centrality (see e.g., Ryan & Hamaker, 2021), which might help identify effective intervention targets for bringing about desired changes in the network structure and dynamics. Further investigation is needed to assess the empirical utility and finite-sample performance of the estimation of these network summary statistics¹³.

In intensive longitudinal studies spanning multiple days, the night gap problem arises due to missing observations be-

¹³The *cTMed* package provides methods for estimating total and indirect effect centrality (Ryan & Hamaker, 2021) and quantifying associated uncertainty using the delta, MC, and PB methods.

tween the last measurement of one day and the first measurement of the next day. We treated this gap as missing data and model dynamics at a single level, assuming continuity during the gap. However, this assumption may not hold for all variables. Future studies can adopt a multilevel approach, nesting observations within days to address the night gap problem.

Autoregressive models often fail to account for persistent trait-like or time-invariant individual differences. The random intercept cross-lagged panel model (RI-CLPM) addresses this limitation by introducing person-specific random intercepts, thereby separating within-person dynamics from stable between-person differences (Hamaker et al., 2015). This separation improves causal interpretation by reducing confounding between trait-level variation and intra-individual processes. Recent developments in CT modeling, such as the continuous-time latent curve model with structured residuals (CT-LCM-SR; Lohmann et al., 2022), offer valuable insights for extending RI-CLPM principles to the CT-VAR framework. These models incorporate random effects for intercepts and slopes while preserving the flexibility of CT dynamics. Leveraging ideas from CT-LCM-SR, future work could explore the integration of random intercepts into CT mediation models, allowing researchers to model both trait-level stability and dynamic mediation processes more accurately.

Lastly, we focused on normally distributed data. However, considering the diversity of data types encountered in substantive research (e.g., categorical, ordinal, and count data), it is essential to explore suitable estimation methods for non-normal data in CT mediation models. Additionally, understanding the implications of using nonnormal data is crucial.

Concluding Remarks

This paper makes several novel contributions to the literature on CT mediation modeling. First, we introduce standardized direct, indirect, and total effects based on steady-state variances—providing scale-free effect size measures for more interpretable and comparable results. Second, we extend the inferential toolkit by evaluating and comparing the delta, MC, and PB methods for estimating SEs and CIs. Through empirical illustrations and simulation studies, we showed that the MC method offers accuracy comparable to PB while being significantly more computationally efficient. These methods, along with automated routines for inference and visualization, are implemented in the open-source R package cTMed, making our contributions both methodologically rigorous and practically accessible.

Open Practices Statement

This study was not preregistered. However, the data and all analysis materials are publicly available on the Open Science Framework (<https://osf.io/qwnmf/>)

and GitHub (<https://github.com/jeksterslab/manCTMed>, <https://jeksterslab.github.io/manCTMed/index.html>). Source code and documentation for the cTMed R package are available on GitHub (<https://github.com/jeksterslab/cTMed>, <https://jeksterslab.github.io/cTMed/index.html>). Additionally, to facilitate reproducibility, we have provided containerized environments that include all necessary dependencies to run the scripts. You can find detailed instructions on setting up and using these containers at: <https://jeksterslab.github.io/manCTMed/articles/containers.html>.

References

- Aalen, O. O., Røysland, K., Gran, J. M., Kouyos, R., & Lange, T. (2016). Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms. *Statistical Methods in Medical Research*, 25(5), 2294–2314. <https://doi.org/10.1177/0962280213520436>
- Aalen, O. O., Røysland, K., Gran, J. M., & Ledgergerber, B. (2012). Causality, mediation and time: A dynamic viewpoint. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 175(4), 831–861. <https://doi.org/10.1111/j.1467-985X.2011.01030.x>
- Arnold, L. (1974). *Stochastic differential equations: Theory and applications*. Wiley.
- Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, 18(2), 265–271. <https://doi.org/10.1214/aoms/1177730442>
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, 40(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5
- Boker, S. M. (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. *Multivariate Behavioral Research*, 37(3), 405–422. https://doi.org/10.1207/s15327906mbr3703_5
- Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, 20, 115. <https://doi.org/10.2307/271084>
- Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2), 144–152. <https://doi.org/10.1111/j.2044-8317.1978.tb00581.x>

- Casella, G., & Berger, R. L. (2002). *Statistical inference*. Thomson Learning.
- Cheung, M. W.-L. (2009a). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, 41(2), 425–438. <https://doi.org/10.3758/brm.41.2.425>
- Cheung, M. W.-L. (2009b). Constructing approximate confidence intervals for parameters with structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(2), 267–294. <https://doi.org/10.1080/10705510902751291>
- Cheung, S. F., & Pesigan, I. J. A. (2023). semlbc: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects, and other derived parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, 30(6), 985–999. <https://doi.org/10.1080/10705512023.2183860>
- Cheung, S. F., Pesigan, I. J. A., & Vong, W. N. (2022). DIY bootstrapping: Getting the nonparametric bootstrap confidence interval in SPSS for any statistics or function of statistics (when this bootstrapping is appropriate). *Behavior Research Methods*, 55(2), 474–490. <https://doi.org/10.3758/s13428-022-01808-5>
- Chow, S.-M., Ho, M.-h. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(2), 303–332. <https://doi.org/10.1080/10705511003661553>
- Chow, S.-M., Losardo, D., Park, J., & Molenaar, P. C. M. (2023). Continuous-time dynamic models: Connections to structural equation models and other discrete-time models. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed.). The Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, 112(4), 558–577. <https://doi.org/10.1037/0021-843X.112.4.558>
- Craig, C. C. (1936). On the frequency function of xy . *The Annals of Mathematical Statistics*, 7(1), 1–15. <https://doi.org/10.1214/aoms/1177732541>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802843>
- de Haan-Rietdijk, S., Voelkle, M. C., Keijser, L., & Hamaker, E. L. (2017). Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01849>
- Deboeck, P. R., & Preacher, K. J. (2015). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(1), 61–75. <https://doi.org/10.1080/10705511.2014.973960>
- Driver, C. C. (2025). Inference with cross-lagged effects—problems in time. *Psychological Methods*, 30(1), 174–202. <https://doi.org/10.1037/met0000665>
- Driver, C. C., Oud, J. H. L., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, 77(5). <https://doi.org/10.18637/jss.v077.i05>
- Driver, C. C., & Voelkle, M. C. (2018). Hierarchical Bayesian continuous time dynamic modeling. *Psychological Methods*, 23(4), 774–799. <https://doi.org/10.1037/met0000168>
- Eddelbuettel, D., & Sanderson, C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics & Data Analysis*, 71, 1054–1063. <https://doi.org/10.1016/j.csda.2013.02.005>
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall. <https://doi.org/10.1201/9780429246593>
- Enders, C. K., Fairchild, A. J., & MacKinnon, D. P. (2013). A Bayesian approach for estimating mediation effects with missing data. *Multivariate Behavioral Research*, 48(3), 340–369. <https://doi.org/10.1080/00273171.2013.784862>
- Fairchild, A. J., & MacKinnon, D. P. (2014). Using mediation and moderation analyses to enhance prevention research. In *Defining prevention science* (pp. 537–555). Springer US. https://doi.org/10.1007/978-1-4899-7424-2_23
- Fairchild, A. J., MacKinnon, D. P., Taborga, M. P., & Taylor, A. B. (2009). R^2 effect-size measures for mediation analysis. *Behavior Research Methods*, 41(2), 486–498. <https://doi.org/10.3758/brm.41.2.486>
- Finkel, S. E. (1995). *Causal analysis with panel data*. Sage.
- Georges, A. R., Alvarez-Bartolo, D., & MacKinnon, D. P. (2023). A sensitivity analysis for temporal bias in cross-sectional mediation. *Psychological Methods*. <https://doi.org/10.1037/met0000628>
- Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, 58(1), 80. <https://doi.org/10.2307/1130293>
- Gollob, H. F., & Reichardt, C. S. (1991). Interpreting and estimating indirect effects assuming time lags really matter. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent*

- advances, unanswered questions, future directions* (pp. 243–259). American Psychological Association. <https://doi.org/10.1037/10099-015>
- Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, 55(292), 708–713. <https://doi.org/10.1080/01621459.1960.10483369>
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424. <https://doi.org/10.2307/1912791>
- Grundy, A. M., Gondoli, D. M., & Blodgett Salafia, E. H. (2007). Marital conflict and preadolescent behavioral competence: Maternal knowledge as a longitudinal mediator. *Journal of Family Psychology*, 21(4), 675–682. <https://doi.org/10.1037/0893-3200.21.4.675>
- 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>
- Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
- Hecht, M., & Zitzmann, S. (2020). A computationally more efficient Bayesian approach for estimating continuous-time models. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(6), 829–840. <https://doi.org/10.1080/10705511.2020.1719107>
- Hecht, M., & Zitzmann, S. (2021). Exploring the unfolding of dynamic effects with continuous-time models: Recommendations concerning statistical power to detect peak cross-lagged effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(6), 894–902. <https://doi.org/10.1080/10705511.2021.1914627>
- Hunter, M. D. (2017). State space modeling in an open source, modular, structural equation modeling environment. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(2), 307–324. <https://doi.org/10.1080/10705511.2017.1369354>
- James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, 69(2), 307–321. <https://doi.org/10.1037/0021-9010.69.2.307>
- Johnson, P. O., & Fay, L. C. (1950). The Johnson-Neyman technique, its theory and application. *Psychometrika*, 15(4), 349–367. <https://doi.org/10.1007/bf02288864>
- Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*, 1, 57–93.
- Jones, J. A., & Waller, N. G. (2015). The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior. *Psychometrika*, 80(2), 365–378. <https://doi.org/10.1007/s11336-013-9380-y>
- Judd, C. M., & Kenny, D. A. (1981). Process analysis. *Evaluation Review*, 5(5), 602–619. <https://doi.org/10.1177/0193841x8100500502>
- Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17(2), 137–152. <https://doi.org/10.1037/a0028086>
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, G. Lindzey, & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., pp. 233–265). McGraw Hill.
- Kreiss, J.-P., & Lahiri, S. N. (2012). Bootstrap methods for time series. In *Time series analysis: Methods and applications* (pp. 3–26). Elsevier. <https://doi.org/10.1016/b978-0-444-53858-1.00001-6>
- Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3). <https://doi.org/10.1214/aos/1176347265>
- Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, 23(2), 244–261. <https://doi.org/10.1037/met0000165>
- Lahiri, P. (2003). On the impact of bootstrap in survey sampling and small-area estimation. *Statistical Science*, 18(2). <https://doi.org/10.1214/ss/1063994975>
- Lohmann, J. F., Zitzmann, S., Voelkle, M. C., & Hecht, M. (2022). A primer on continuous-time modeling in educational research: An exemplary application of a continuous-time latent curve model with structured residuals (CT-LCM-SR) to PISA data. *Large-scale Assessments in Education*, 10(1). <https://doi.org/10.1186/s40536-022-00126-8>
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-27752-1>
- Lyapunov, A. M. (1992). The general problem of the stability of motion. *International Journal of Control*, 55(3), 531–534. <https://doi.org/10.1080/00207179208934253>
- MacKinnon, D. P. (1994). Analysis of mediating variables in prevention and intervention research. *NIDA research monograph*, 139, 127–153.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Erlbaum Psych Press. <https://doi.org/10.4324/9780203809556>

- Mackinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17*(2), 144–158. <https://doi.org/10.1177/0193841x9301700202>
- MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods, 39*(3), 384–389. <https://doi.org/10.3758/bf03193007>
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173–181. <https://doi.org/10.1023/a:1026595011371>
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23–44. <https://doi.org/10.1037/1082-989x.12.1.23>
- Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research, 46*(5), 816–841. <https://doi.org/10.1080/00273171.2011.606716>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide. Eighth edition*. Los Angeles, CA, Muthén & Muthén.
- Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2015). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika, 81*(2), 535–549. <https://doi.org/10.1007/s11336-014-9435-8>
- Oehlert, G. W. (1992). A note on the delta method. *The American Statistician, 46*(1), 27–29. <https://doi.org/10.1080/00031305.1992.10475842>
- Oh, H., Hunter, M. D., & Chow, S.-M. (2025). Measurement model misspecification in dynamic structural equation models: Power, reliability, and other considerations. *Structural Equation Modeling: A Multidisciplinary Journal, 32*(3), 511–528. <https://doi.org/10.1080/10705511.2025.2452884>
- O'Laughlin, K. D., Martin, M. J., & Ferrer, E. (2018). Cross-sectional analysis of longitudinal mediation processes. *Multivariate Behavioral Research, 53*(3), 375–402. <https://doi.org/10.1080/00273171.2018.1454822>
- Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods, 16*(4), 468–490. <https://doi.org/10.1037/a0024375>
- Oravecz, Z., Wood, J., & Ram, N. (2018). On fitting a continuous-time stochastic process model in the Bayesian framework. In *Continuous time modeling in the behavioral and related sciences* (pp. 55–78). Springer International Publishing. https://doi.org/10.1007/978-3-319-77219-6_3
- O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs, 79*(2), 171–181. <https://doi.org/10.15288/jsad.2018.79.171>
- O'Rourke, H. P., & MacKinnon, D. P. (2019). The importance of mediation analysis in substance-use prevention. In *Advances in prevention science* (pp. 233–246). Springer International Publishing. https://doi.org/10.1007/978-3-030-00627-3_15
- Orzek, J. H., & Voelkle, M. C. (2023). Regularized continuous time structural equation models: A network perspective. *Psychological Methods, 28*(6), 1286–1320. <https://doi.org/10.1037/met0000550>
- Ou, L., Hunter, M. D., & Chow, S.-M. (2019). What's for dynr: A package for linear and nonlinear dynamic modeling in R. *The R Journal, 11*(1), 91. <https://doi.org/10.32614/rj-2019-012>
- Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-11760-2_7
- Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*(2), 199–215. <https://doi.org/10.1007/bf02294374>
- Pawitan, Y. (2013). *In all likelihood: Statistical modelling and inference using likelihood*. Oxford University Press.
- Pesigan, I. J. A., & Cheung, S. F. (2020). SEM-based methods to form confidence intervals for indirect effect: Still applicable given nonnormality, under certain conditions. *Frontiers in Psychology, 11*. <https://doi.org/10.3389/fpsyg.2020.571928>
- Pesigan, I. J. A., & Cheung, S. F. (2023). Monte Carlo confidence intervals for the indirect effect with missing data. *Behavior Research Methods, 56*(3), 1678–1696. <https://doi.org/10.3758/s13428-023-02114-4>
- Pesigan, I. J. A., Sun, R. W., & Cheung, S. F. (2023). betaDelta and betaSandwich: Confidence intervals for standardized regression coefficients in R. *Multivariate Behavioral Research, 58*(1), 1–20. <https://doi.org/10.1080/00273171.2023.2290001>

- ate Behavioral Research*, 58(6), 1183–1186. <https://doi.org/10.1080/00273171.2023.2201277>
- Politis, D. N. (2003). The impact of bootstrap methods on time series analysis. *Statistical Science*, 18(2). <https://doi.org/10.1214/ss/1063994977>
- Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313. <https://doi.org/10.1080/01621459.1994.10476870>
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31(4), 437–448. <https://doi.org/10.3102/10769986031004437>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. <https://doi.org/10.3758/brm.40.3.879>
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93–115. <https://doi.org/10.1037/a0022658>
- Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, 6(2), 77–98. <https://doi.org/10.1080/19312458.2012.679848>
- R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Rogosa, D. R. (1979). Causal models in longitudinal research: Rationale, formulation, and interpretation. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal methodology in the study of behavior and development*. Academic Press.
- Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88(2), 245–258. <https://doi.org/10.1037/0033-2909.88.2.245>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>
- Ryan, O., & Hamaker, E. L. (2021). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*, 87(1), 214–252. <https://doi.org/10.1007/s11336-021-09767-0>
- Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous-time approach to intensive longitudinal data: What, why, and how? In *Continuous time modeling in the behavioral and related sciences* (pp. 27–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-77219-6_2
- 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(2), 206–221. <https://doi.org/10.1037/met0000062>
- Schuurman, N. K., & Hamaker, E. L. (2019). Measurement error and person-specific reliability in multilevel autoregressive modeling. *Psychological Methods*, 24(1), 70–91. <https://doi.org/10.1037/met0000188>
- Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01038>
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422–445. <https://doi.org/10.1037/1082-989x.7.4.422>
- Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-52452-8>
- Singer, H. (2012). SEM modeling with singular moment matrices part II: ML-estimation of sampled stochastic differential equations. *The Journal of Mathematical Sociology*, 36(1), 22–43. <https://doi.org/10.1080/0022250x.2010.532259>
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290. <https://doi.org/10.2307/270723>
- Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological Methodology*, 16, 159. <https://doi.org/10.2307/270922>
- Sobel, M. E. (1987). Direct and indirect effects in linear structural equation models. *Sociological Methods & Research*, 16(1), 155–176. <https://doi.org/10.1177/0049124187016001006>
- Staudenmayer, J., & Buonaccorsi, J. P. (2005). Measurement error in linear autoregressive models. *Journal of the American Statistical Association*, 100(471), 841–852. <https://doi.org/10.1198/016214504000001871>
- Stoffer, D. S., & Wall, K. D. (1991). Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. *Journal of the American Statistical Association*, 86(416), 1024–1033. <https://doi.org/10.1080/01621459.1991.10475148>
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288.

- <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(3), 273–282. <https://doi.org/10.1111/j.1467-9868.2011.00771.x>
- Tofghi, D., & Kelley, K. (2019). Indirect effects in sequential mediation models: Evaluating methods for hypothesis testing and confidence interval formation. *Multivariate Behavioral Research*, 55(2), 188–210. <https://doi.org/10.1080/00273171.2019.1618545>
- Tofghi, D., & Kelley, K. (2020). Improved inference in mediation analysis: Introducing the model-based constrained optimization procedure. *Psychological Methods*, 25, 496–515. <https://doi.org/10.1037/met0000259>
- Tofghi, D., & MacKinnon, D. P. (2015). Monte Carlo confidence intervals for complex functions of indirect effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(2), 194–205. <https://doi.org/10.1080/10705511.2015.1057284>
- Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the brownian motion. *Physical Review*, 36(5), 823–841. <https://doi.org/10.1103/physrev.36.823>
- Usami, S. (2020). On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: Interpretation of cross-lagged parameters and model choice. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(3), 331–344. <https://doi.org/10.1080/10705511.2020.1821690>
- 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>
- van Montfort, K., Oud, J. H. L., & Voelkle, M. C. (Eds.). (2018). *Continuous time modeling in the behavioral and related sciences*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-77219-6>
- Venzon, D. J., & Moolgavkar, S. H. (1988). A method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, 37(1), 87. <https://doi.org/10.2307/2347496>
- Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician*, 66(2), 124–127. <https://doi.org/10.1080/00031305.2012.687494>
- Voelkle, M. C., & Oud, J. H. L. (2012). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, 66(1), 103–126. <https://doi.org/10.1111/j.2044-8317.2012.02043.x>
- 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(2), 176–192. <https://doi.org/10.1037/a0027543>
- Wang, L., & Zhang, Q. (2020). Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs. *Psychological Methods*, 25(3), 271–291. <https://doi.org/10.1037/met0000235>
- Yuan, K.-H., & Chan, W. (2011). Biases and standard errors of standardized regression coefficients. *Psychometrika*, 76(4), 670–690. <https://doi.org/10.1007/s11336-011-9224-6>
- Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14(4), 301–322. <https://doi.org/10.1037/a0016972>
- Zhang, G. (2018). Testing process factor analysis models using the parametric bootstrap. *Multivariate Behavioral Research*, 53(2), 219–230. <https://doi.org/10.1080/00273171.2017.1415123>
- Zhang, G., & Browne, M. W. (2010). Bootstrap standard error estimates in dynamic factor analysis. *Multivariate Behavioral Research*, 45(3), 453–482. <https://doi.org/10.1080/00273171.2010.483375>
- Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., Hamaker, E. L., Sham-sollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2019). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods*, 23(4), 651–687. <https://doi.org/10.1177/1094428119847278>
- Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., Hamaker, E. L., Sham-sollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2019). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*, 23(4), 688–716. <https://doi.org/10.1177/1094428119847280>