

Multilevel Latent Differential Structural Equation Model with Short Time Series and Time-Varying Covariates: A Comparison of Frequentist and Bayesian Estimators

Young Won Cho^a , Sy-Miin Chow^{a,b}, Christina M. Marini^c, and Lynn M. Martire^{a,d}

^aDepartment of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, USA; ^bSocial Science Research Institute, The Pennsylvania State University, University Park, PA, USA; ^cDepartment of Psychology, Gordon F. Derner School of Psychology, Adelphi University, Garden City, NY, USA; ^dThe Center for Healthy Aging, The Pennsylvania State University, University Park, PA, USA

ABSTRACT

Continuous-time modeling using differential equations is a promising technique to model change processes with longitudinal data. Among ways to fit this model, the Latent Differential Structural Equation Modeling (LDSEM) approach defines latent derivative variables within a structural equation modeling (SEM) framework, thereby allowing researchers to leverage advantages of the SEM framework for model building, estimation, inference, and comparison purposes. Still, a few issues remain unresolved, including performance of multilevel variations of the LDSEM under short time lengths (e.g., 14 time points), particularly when coupled multivariate processes and time-varying covariates are involved. Additionally, the possibility of using Bayesian estimation to facilitate the estimation of multilevel LDSEM (M-LDSEM) models with complex and higher-dimensional random effect structures has not been investigated. We present a series of Monte Carlo simulations to evaluate three possible approaches to fitting M-LDSEM, including: frequentist single-level and two-level robust estimators and Bayesian two-level estimator. Our findings suggested that the Bayesian approach outperformed other frequentist approaches. The effects of time-varying covariates are well recovered, and coupling parameters are the least biased especially using higher-order derivative information with the Bayesian estimator. Finally, an empirical example is provided to show the applicability of the approach.

KEYWORDS

Latent differential equations; Bayesian statistics; coupled damped linear oscillators; affect dynamics

Differential equation models quantify change processes continuously, thereby providing projection of a system's values at any arbitrary time point, even in situations where the observed time intervals are irregularly spaced. Latent Differential Structural Equation Models (LDSEM), proposed by Boker et al. (2004), is a framework for fitting differential equation models as structural equation models (SEMs) by specifying derivatives as latent variables through constrained loadings that are similar to those used to define intercepts, slopes, and higher-order slopes in growth curve models. In doing so, LDSEM allows researchers to leverage the benefits afforded by the SEM framework, such as easy inclusion of antecedent and outcome measures, and direct consolidation of multiple measures, possibly measured on mixed (e.g., continuous, ordinal, and nominal) scales. In addition, LDSEM shows notable improvements in

performance compared to alternative two-stage model fitting approaches where derivatives are first computed by means of some approximation methods and subsequently used as observed variables in SEM (Boker et al., 2010; Boker & Graham, 1998; Chow, 2019).

There is no scarcity of applications utilizing differential equation models in the social and behavioral sciences. For example, the linear damped oscillator model, a model used to describe the movements of a swinging pendulum under friction, has been used to describe how individuals regulate their emotions (Chow et al., 2005; Katinka et al., 2020). Bivariate extensions of the damped oscillator model, which feature two swinging pendulums as coupled to one another, have been used to examine the relationships between emotional eating and hormone change (Hu et al. (2014), interpersonal co-regulation in body movement, emotions, and

physiological signals (Chow et al., 2010; Ferrer & Helm, 2013; Feinberg et al., 2017).

Unresolved methodological issues

Several unresolved issues need to be explored to bridge the gaps between the current LDSEM models and the needs of practical researchers. These issues, elaborated in turn below, were crucial for testing key aspects of our motivating empirical example, and formed the basis of the methodological extensions proposed and investigated in the present article.

Multilevel differential equation models and Bayesian inference

Parameter estimation for differential equation models, especially those involving multilevel data and high-dimensional random effect structures, is not a trivial task. Frequentist estimation approaches typically require integration—whether analytically, numerically (Durham & Gallant, 2002), or *via* alternative Monte Carlo approaches (Chow et al., 2016), over the random-effects distributions. Since most of these integrals lack closed-form solutions, the model estimation process is computationally expensive. Even though significant advances have been achieved in fitting multilevel latent variable models using frequentist approaches (Asparouhov & Muthén, 2007), including techniques for computing robust standard errors (Huber, 1967; White, 1980, 1994), estimation results are often unsatisfactory in situations involving a larger number of and/or complex random effect structures. In contrast, in the Bayesian paradigm, estimation of all parameters, latent variables, and random effects is performed *via* Markov Chain Monte Carlo (MCMC) techniques that can handle integration over much higher dimensions with substantially fewer distributional restrictions. Therefore, using a Bayesian framework to conduct statistical inference offers clear pragmatic advantages (Chen et al., 2023; Elerian et al., 2001; Lu et al., 2015, 2019; Mbalawata et al., 2013; Oravecz et al., 2011). Thus, one of key objectives of this study is to compare the performance of Bayesian and robust frequentist approaches for fitting multilevel LDSEM (M-LDSEM) models to multilevel data. The descriptions of the robust frequentist and Bayesian estimators used in the present simulation study are presented later in the Modeling Approach section.

Finite-length “intensive” longitudinal data

Another notable gap between published simulation findings involving the LDSEM and applications that

are “real-world” resides in the number of repeated measures available from participants. According to review papers about ecological momentary assessment (EMA) designs as adopted in the studies of mood, sleep, stress, social interaction, and clinical outcomes (Konjarski et al., 2018; Liu et al., 2019; Rodríguez-Blanco et al., 2018; Yang et al., 2019), the number of occasions typically ranges from 3 to 336 (median = 14, mean = 20.1), with 14 days being one of the most popular sample size configurations because it provides the minimum number of repeated measures needed to capture some recurrence in weekly dynamics, in addition to revealing other momentary dynamics that unfold on a faster (e.g., daily) time scale.

Although some EMA studies do incorporate multiple measurements a day, such sampling density may not be feasible for many other constructs of interest (e.g., sleep quality). Consequently, it is common for longitudinal studies in the behavioral sciences to comprise relatively few measurement occasions compared to data from the physical sciences, engineering, and economics, in which applications of differential equations are prevalent. Even though a few recent studies did provide some pointers on the recommended sample size configurations (in terms of the numbers of participants and time points) for fitting multilevel dynamic models, most of this work focuses heavily on discrete-time dynamic models, not continuous-time models (Li et al., 2022; Schultzberg & Muthén, 2018). Thus, we aim to evaluate the performance of the LDSEM under commonly adopted EMA sample size configurations, and possible methodological enhancements that aid estimation under such data constraints.

Time-varying covariates

Including exogenous time-varying covariates in LDSEM allows researchers to address a broader range of questions beyond just exploring the system’s internal dynamics. For instance, in our motivating example, we are interested in examining the implications of providing social support (a time-varying covariate) to a partner with a chronic health condition on a couple’s ebb and flow in positive emotion.

Incorporating time-varying covariates into LDSEM is meaningful in terms of the modeling aspect as well as the research question. Figure 1 shows how the trajectory of an oscillator differs in the absence of and with the influence of a simulated time-varying covariate. If a time-varying covariate is omitted from a model when it should be included, as in the second plot in Figure 1, estimates of the unconditional models do not fully reflect the time-varying nature of the

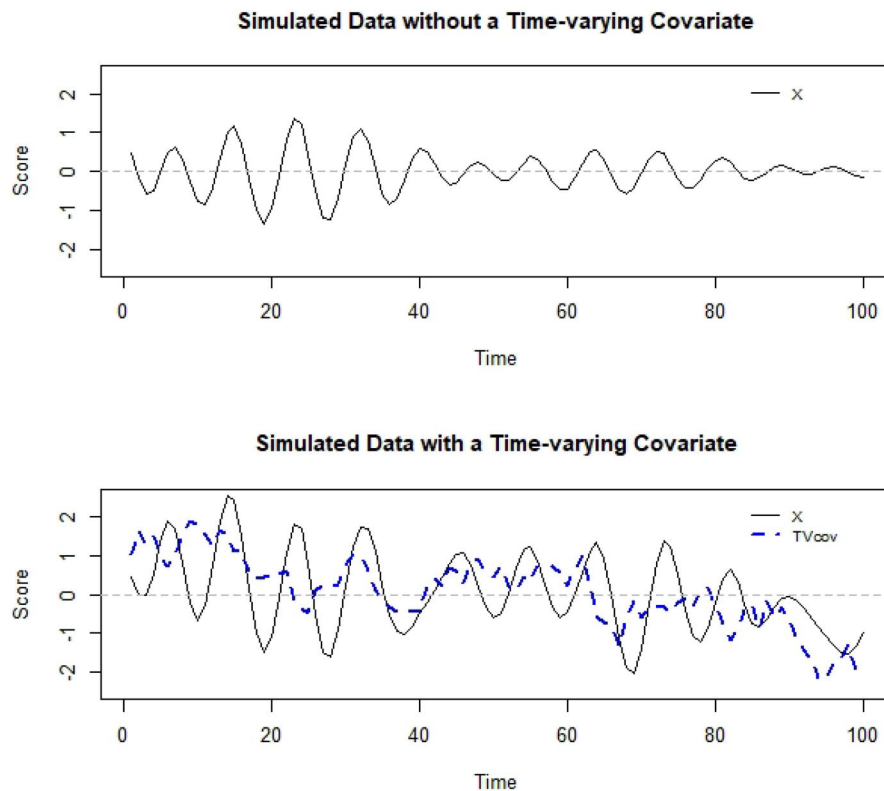


Figure 1. Simulated trajectories with and without the effect of a time-varying covariate.

Note. This figure demonstrates one possible way of conceptualizing the effects of a time-varying covariate, as adopted in our motivating empirical example. In this case, the levels of the time-varying covariate are hypothesized to influence the second derivatives (i.e., accelerations/decelerations, or changes and curvatures in the rises and declines) of the process of interest (X). The inclusion of this time-varying covariate has the effect of shifting the equilibrium (the value to which the oscillator settles in the long run) vertically upward/downward depending on the values of the covariate.

data, and may lead to substantively misguided conclusions. For example, when an oscillatory model was fitted to test data with a true frequency parameter of -0.5 , the estimate in the model with time-varying covariates was -0.497 , while it was -0.239 without estimating the effect of the time-varying covariates. Thus, in this article, we seek to examine a possible way of incorporating time-varying covariate effects into the LDSEM framework, key methodological considerations, from an estimation standpoint.

LDSEM with coupled dynamic processes and recent methodological enhancements

Even though bivariate variations of the linear oscillator model are gaining traction in studies of dyadic and interpersonal dynamics (Niven, 2017), few Monte Carlo simulation studies were available to clarify the performance of the LDSEM method in fitting coupled differential equation models. The one exception that existed indicated that the coupling parameters that summarize the mutual influences of two dyad members on each other were not accurately estimated under the LDSEM approach (Hu et al., 2014). Thus, a

comparison of the currently available estimation methods for recovering coupling-related parameters, especially under M-LDSEM, is warranted.

We consider in the present paper two possible enhancements to current estimation methods for fitting M-LDSEM. One of them is to use robust variations of frequentist (maximum likelihood) estimation (Asparouhov, 2005; Freedman, 2006) to circumvent a known issue in use of time-delay embedded data to fit LDSEM their discrete-time counterparts (Molenaar, 1985). Previous simulations have shown that when time-delay embedded data were used, the standard error estimates tended to be biased, despite relatively satisfactory point estimates (Chow et al., 2010). To this end, robust maximum likelihood approaches offer corrections to the standard error estimates and corresponding violations of the independence assumptions of the residuals, and may thus offer some improvements in estimation results compared to standard frequentist approaches. In addition, previous comparisons of Bayesian and frequentist approaches to estimating dynamic and longitudinal models also indicated some advantages of the former over the

latter, especially under finite sample sizes (Zhang & Nesselroade, 2007; Zhang et al., 2007, 2008).

In summary, this paper aims to build a differential equation model that echoes substantive needs and then evaluate whether variations of the LDSEM approaches, which incorporate recent advanced modeling techniques, perform well to fit the model. To do so, we constructed a simulation model from our motivating example: a coupled damped oscillator model with mixed effects and time-varying covariates. Then, we ran a Monte Carlo simulation to examine the performance of the different LDSEM methods. Specifically, we compared (1) standard LDSEM methods versus noble LDSEM methods that leverage higher-order derivative information and (2) robust frequentist versus Bayesian approaches. We also examined whether the length of the time series limits the use of the LDSEM technique.

The rest of the article is organized as follows. First, we introduce our motivating example and its corresponding differential equation model. Second, backgrounds for the LDSEM and its fourth-order variation (Boker et al., 2004) are briefly reviewed. Then, a Monte Carlo simulation is conducted to evaluate the performance of each LDSEM variation. The recommended LDSEM procedure is applied to empirical data based on the simulation result.

Motivating example

For married individuals suffering from chronic illness, their spouses often become primary caregivers (Sanders & Power, 2009), and such spousal support for patients is related to couples' well-being. Our modeling framework was motivated by data from 22-day daily diary data from patients with knee osteoarthritis (OA) and their spouses (Martire et al., 2013). Both patients and spouses reported daily ratings of positive affect and spousal instrumental support in coping with OA. Spousal instrumental support means tangible aid from spouses, such as housework and personal care, to help patients avoid pain. Since spouses diagnosed with moderate to severe OA were excluded from participation, each couple member evaluated support provision only from spouses. We sought to study couples' emotional dynamics and how instrumental support from spouses for patients influenced couples' emotional regulation processes. Specifically, we were interested in addressing the extent and directionality of the patients' and their spouses' influences on each other (i.e., emotional covariation). In this section, we begin by describing a

univariate version of our proposed empirical model before transitioning to its bivariate variation motivated by our empirical example.

Damped linear oscillator model as a model of individuals' emotion regulation

Researchers differ slightly on the exact functional forms of such emotion regulation processes. Some conceptualized the corresponding trajectories as monotonic (i.e., without changes in directions) returns to baseline (Koval et al., 2012; Kuppens et al., 2010; Oravecz et al., 2011), which can be operationalized as a first-order differential equation focusing on representing the amounts of instantaneous changes (first derivatives) in affect. Other researchers have used the analogy of emotion as a "thermostat" to describe the homeostatic process through which individuals self-regulate their emotions toward some affective (possibly person-specific) baseline (Chow et al., 2005; Larsen, 2000) in the absence of interventions or actions from other individuals. Mathematically, this may be captured using a second-order differential equation, which delineates changes in second derivatives, or the amounts of instantaneous changes from one time interval to the next (e.g., accelerations and decelerations). Under this framework, individuals' self-regulatory processes are typically viewed as oscillatory in nature, even though monotonic return to baseline (i.e., approaching baseline without changes in directions of such changes) is possible in some individuals. In other words, individuals are conceptualized to show some ebbs and flows—sometimes "over-" or even "under-regulating" to levels beyond or less than their affective baselines—before they settle into their baselines.

The second-order differential equation used by Chow et al. (2005) to depict daily ebbs and flows of emotions as a "thermostat", often denoted in the literature as a damped linear oscillator model (Boker & Nesselroade, 2002), is defined as:

$$\ddot{x}_{it} = \eta x_{it} + \zeta \dot{x}_{it} \quad (1)$$

where x_{it} is the level of positive affect relative to its typical level (for person i at time t); $\dot{x}_{it}$ and $\ddot{x}_{it}$ are the first derivative and the second derivative of positive affect. The first derivative expresses the rate of change, or the amount of change in x_{it} from one time point to the next as the time interval gets infinitely small. The second derivative expresses how rapidly the first derivative is changing (changes in the rates of change), indicating whether the instantaneous changes in x_{it} are increasing (accelerating) or decreasing

(decelerating). η is a parameter that governs the frequency, or how rapidly x_{it} changes over time, and ζ is a parameter that controls the changes in amplitude of the process over time, with negative values of ζ leading to damping (decrease in amplitude) and positive values to amplification (increase in amplitude) of the oscillations over time. In specific ranges of η and ζ (Zill, 1993), the process shows oscillations with damping over time (see Figure 2 for examples).

Coupled damped linear oscillator model as a representation of couples' dynamics

Interpersonal emotion regulation studies (Elfenbein, 2014; Niven, 2017) suggest that there may be reciprocal influences, or coupling, between patients and their caregivers in everyday emotion regulation. There can be various patterns of co-regulation or mutual influence. For example, a patient's affect may easily impact the state of a spouse but may not be influenced by the spouse's state. In close relationships, when their partner is highly aroused, one may become calm gradually, while another may become aroused together.

Consistent with evidence that emerged from other empirical work involving EMA data of affect in dyads (Feinberg et al., 2017), we sought to use a coupled damped linear oscillator model as the basis of our modeling framework to examine self-regulation as well as co-regulation—the extent to which two

individuals (patients and spouses in this case) influence each other in their emotion regulation processes. Mathematically, a coupled damped linear oscillator model is a bivariate version of the model in Equation (1) and is expressed as follows:

$$\ddot{x}_{it} = \eta_1 x_{it} + \zeta_1 \dot{x}_{it} + \gamma_1 y_{it} \quad (2)$$

$$\ddot{y}_{it} = \eta_2 y_{it} + \zeta_2 \dot{y}_{it} + \gamma_2 x_{it}, \quad (3)$$

in which x refers to the positive affect of the patient and y is the positive affect of the spouse in our example. Equation (2) represents the equation of the patient, and Equation (3) is for the spouse. In addition to the parameters η_1 , η_2 , ζ_1 and ζ_2 , this model includes two coupling parameters, γ_1 and γ_2 . γ_1 represents the degree to which a patient's emotion regulation process is influenced by the current level of the spouse's affect (y_{it}) in addition to the patient's own affect level (x_{it}) and rate of change ($\dot{x}_{it}$). Hence, the coupling parameter can also be interpreted as permeability or sensitivity (Butler, 2011). For example, a high absolute value of the coupling parameter can indicate that the person has more permeable interpersonal boundaries, which means that the partner's affect can easily impact the state of the other. For details on the interpretation of coupling parameters, see Hu et al. (2014).

This model is increasingly employed because of its ability to capture the extent and directionality of the co-regulation dynamics (e.g., which partner is leading the covariations in the dyad's dynamics as a whole), which cannot be tested by examining the day-to-day covariance of the variable. For instance, Steele and Ferrer (2011) found that, on average across the sample, both females and males were sensitive to their partners' overall affect levels but only females were sensitive to males' rate of change of positive affect.

Instrumental support as time-varying covariate and multilevel extensions

A second key research question was the role of instrumental support on both patients' and spouses' emotion regulation processes. The effect of spousal support has been reported as being rather mixed. Reduced mental and physical health of caregivers has been linked to caring for a chronically sick spouse in daily activities (Monin & Schulz, 2009); however, it has also been found that caring for a partner provides advantages for caregivers' mental health (Beach et al., 2000; Poulin et al., 2010). Likewise, receiving instrumental support from spouses can benefit patients' well-being, while it does not reduce patients'

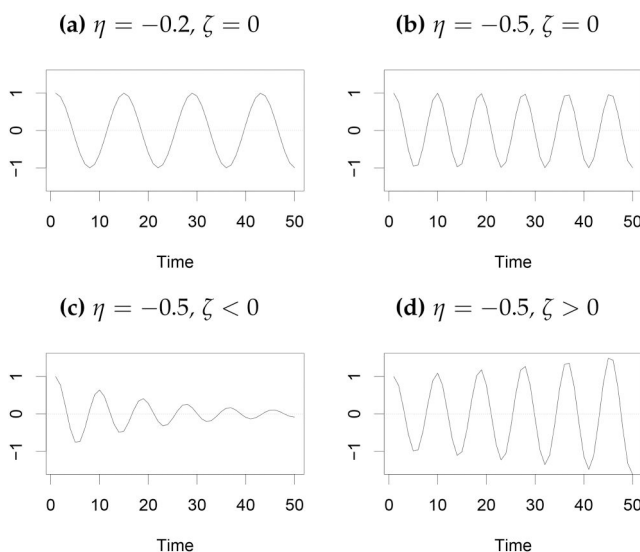


Figure 2. Simulated trajectories according to η and ζ .
Note. When a damping parameter $\zeta = 0$, the amplitude of the process does not change over time unless external forces work on it, and faster oscillation occurs under a more negative value of η (see subfigures A and B). If $\zeta < 0$, the process shows damping over time, and if $\zeta > 0$, the amplitude of the oscillation will increase over time (see subfigures C and D).

depressive symptoms and can backfire when it emphasizes the inability of the recipient (De Leeuw et al., 2000; Reinhardt et al., 2006).

Such inconsistency may be resolved by considering inter-individual differences in the effect of daily instrumental support provision/receipt. For instance, providing a high level of support on a day is something that happens in an extraordinary situation to someone who does not usually take care of patients, and providing higher support than their usual level may greatly impact their emotions of the day. Conversely, based on the desensitization theory (McGlynn et al., 1981), for people whose typical level of support provision is high, since caregiving behavior is just an ordinary event that usually happens, their response to the elevated daily support provision would be reduced. To incorporate such interindividual differences, we will use a multilevel model that allows the effect of daily support to differ between couples. Consequently, our motivating model is expressed as:

$$\begin{aligned} \ddot{Patient}_{it} &= \eta_p Patient_{it} + \gamma_{pi} Spouse_{it} + \beta_{pi} PatientSupport_{it} + e_{p,it} \\ \gamma_{pi} &= \gamma_{p0} + \gamma_{p1} PatientSupport_{it} + u_{\gamma,pi} \\ \beta_{pi} &= \beta_{p0} + \beta_{p1} PatientSupport_{it} + u_{\beta,pi} \\ \ddot{Spouse}_{it} &= \eta_s Spouse_{it} + \gamma_{si} Patient_{it} + \beta_{si} SpouseSupport_{it} + e_{s,it} \\ \gamma_{si} &= \gamma_{s0} + \gamma_{s1} PatientSupport_{it} + u_{\gamma,si} \\ \beta_{si} &= \beta_{s0} + \beta_{s1} SpouseSupport_{it} + u_{\beta,si} \end{aligned} \quad (4)$$

where $Patient_{it}$ and $Spouse_{it}$ refer to positive affect of each couple member. Spousal instrumental support that patients or spouses reported ($PatientSupport_{it}$, $SpouseSupport_{it}$) is a time-varying covariate in the patient's or spouse's equation, respectively. $PatientSupport_{it}$ and $SpouseSupport_{it}$ represent time-invariant covariates, the average levels of instrumental support over 22 days.

Using this model, we can examine not only whether the daily positive affect of a couple shows an oscillatory self-regulation process (η_p and η_s) and whether their affect dynamics co-regulate each other (coupling effect; γ_{p0} and γ_{s0}), but also whether the daily spousal instrumental support, the time-varying covariate, is associated with the accelerations and decelerations in the daily affect of each couple member (β_{p0} and β_{s0}). Further, to capture possible sources of interindividual differences, the effect of daily instrumental support is allowed to differ between couples, and time-invariant covariates are included as predictors in the model. As co-regulation patterns could also vary across couples, random variances for coupling effects and time-invariant covariates to explain the random variance are included in the model.

We described our motivating example and introduced univariate and bivariate dynamic models that consider the over-time evolution of each dyad member's emotions. We also illustrated a model that shows how time-varying and time-invariant covariates can be incorporated into those models. We will revisit this example in section Empirical Example by fitting the model in Equation (4) to the empirical data.

Steps for fitting M-LDSEM

In this section, we outline the procedures involved in fitting an M-LDSEM into two major steps: (1) preparations of time-delay embedded data, and (2) specification of M-LDSEM, including possible ways to leverage higher-derivative information from the dependent variables and time-varying covariates.

Step 1: Preparations of time-delay embedded data

To fit differential equation models to empirical data using the LDSEM method in SEM software, the first step is to prepare the data that enables specification of derivatives as latent factors (Boker et al., 2004). Time delay embedding is a data preparation process that involves rearranging individuals' raw time series data to create successively lagged (or delayed) blocks of the original time series. For example, consider a 100×1 time series consisting of data of an individual measured on one variable over 100 time points. To construct time-delay embedded data, a researcher first has to determine the number of embedding dimensions (d), namely, how many lagged occasions have to be paired with the data at each time point. With $d=6$, the value of x_t at time t (lag of 0) would be paired with the corresponding measurements from that person at lag of 1, 2, ..., 5 (x_{t-5} , x_{t-4} , x_{t-3} , x_{t-2} , x_{t-1} , x_t), thus creating a total of six variables that represent concurrent and lagged versions of the original time series. In this case, the final input data matrix for model fitting, denoted as $\mathbf{X}^{(6)}$, is a 95×6 data matrix, with structure shown as follows.

Original time series:

$$\mathbf{x} = [x_{1,1} \quad x_{1,2} \quad x_{1,3} \quad \dots \quad x_{1,98} \quad x_{1,99} \quad x_{1,100}]$$

Time-delay embedded data of $d = 6$:

$$\mathbf{X}^{(6)} = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & x_{1,5} & x_{1,6} \\ x_{1,2} & x_{1,3} & x_{1,4} & x_{1,5} & x_{1,6} & x_{1,7} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,95} & x_{1,96} & x_{1,97} & x_{1,98} & x_{1,99} & x_{1,100} \end{bmatrix}$$

where $x_{i,t}$ is the value of the variable x for the i th

individual at time t . von Oertzen and Boker (2010) has shown that this time delay embedding technique improved the estimation properties of parameters for models of intraindividual dynamic processes.

Step 2: Specification of M-LDSEM as factor model with constrained loadings

After time-delay embedding, a differential equation model is fitted to the time-delay embedded data. In the LDSEM method (Boker et al., 2004), the derivatives (i.e., x , $\dot{x}$, $\ddot{x}$) are specified as latent factors in an SEM framework with specialized loadings. This is what differentiates the LDSEM method from previous two-step approaches such as the Generalized Local Linear Approximation (GLLA) method, in which derivatives are first calculated explicitly and then used as observed variables in differential equation models (Boker et al., 2010). Additionally, GLLA assumes equal intervals between measurement occasions, which may not always be the case in LDSEM.

Specifically, the LDSEM method simultaneously constructs the latent derivative factors in the measurement model and estimates the parameters of the differential equation (i.e., the relationships between the latent derivative variables) in the structural model. The first diagram in Figure 3 represents a standard 2nd-order differential equation model using the LDSEM approach. In the measurement model, the latent derivatives are constructed by specifying predetermined factor loadings ($\mathbf{L}$) for the relationships between the latent derivatives (factors) and the observed measures (indicators). The six indicators in Figure 3 correspond to columns in the time-delay embedded matrix (e.g., $\mathbf{X}^{(6)}$). Factor loading, $\mathbf{L}$, is determined by polynomials of the Taylor series expansion (Chow et al., 2016) based on the time-delay embedding dimension (d), the designated time interval (Δt), and the highest order of the derivatives used. Figure 3 shows an example of $\mathbf{L}$ with $d=6$, $\Delta t=1$, and the highest order = 2. For more details on the calculation of $\mathbf{L}$, see Boker et al. (2010). This particular specification of LDSEM can be viewed as an SEM consisting of a factor model with specialized constrained loadings, and is also comparable to how fixed loadings are used to construct the latent intercept and slopes in a latent growth curve model in the SEM framework.

Extension from single-level to multilevel LDSEM

We introduce M-LDSEM by integrating LDSEMs with multilevel modeling within the framework of

multilevel structural equation models (MSEM). MSEM is employed when dealing with nested data structures, necessitating the simultaneous modeling of latent variables, measurement errors, and pathways. In the subsequent sections, we will provide an overview of MSEM and then present the M-LDSEM.

Multilevel structural equation modeling

In MSEM, observed variables can be decomposed into components that vary within groups (in this context, within individuals across days) and between groups (across individuals). For instance, with daily positive emotion scores collected over T days ($t = 1, \dots, T$) for N individuals ($i = 1, \dots, N$), the daily observations are nested within individuals, creating a multilevel structure. The observed individual-level scores can be decomposed into within and between components as $\mathbf{x}_{it} = \mathbf{x}_{wit} + \mathbf{x}_{bi}$, where $\mathbf{x}_{wit}$ reflects daily fluctuations in positive emotions for individual i , and $\mathbf{x}_{bi}$ captures consistent individual differences over time.

Each component is modeled as follows:

$$\mathbf{x}_{wit} = \mathbf{\Lambda}_W \boldsymbol{\eta}_W + \boldsymbol{\varepsilon}_W \quad (5)$$

$$\mathbf{x}_{bi} = \boldsymbol{\mu} + \mathbf{\Lambda}_B \boldsymbol{\eta}_B + \boldsymbol{\varepsilon}_B \quad (6)$$

where $\boldsymbol{\eta}_W$ and $\boldsymbol{\eta}_B$ are within- and between-level latent factors; $\mathbf{\Lambda}_W$ and $\mathbf{\Lambda}_B$ are within- and between-level factor loadings; $\boldsymbol{\mu}$ is the overall mean across individuals; $\boldsymbol{\varepsilon}_W$ and $\boldsymbol{\varepsilon}_B$ are within- and between-level residuals. Time-varying covariates would need to be inserted into $\boldsymbol{\eta}_W$ in Equation (5).

The structural equations for latent variables can be expressed as:

$$\boldsymbol{\eta}_W = \boldsymbol{\Gamma}_W \boldsymbol{\eta}_W + \mathbf{e}_W \quad (7)$$

$$\boldsymbol{\eta}_B = \boldsymbol{\Gamma}_B \boldsymbol{\eta}_B + \mathbf{e}_B \quad (8)$$

where $\boldsymbol{\Gamma}_W$ and $\boldsymbol{\Gamma}_B$ contain coefficients that define structural relationships between factors; $\mathbf{e}_W$ and $\mathbf{e}_B$ contain residual terms.

Multilevel LDSEM

In a similar way, the specific M-LDSEM that allows for individual-specific dynamic coefficients with one between-level covariate is stated as follows:

$$\mathbf{x}_{it}^{(6)} = \mathbf{x}_{wit}^{(6)} + \mathbf{x}_{bi}^{(6)} \quad (9)$$

where $\mathbf{x}_{it}^{(6)}$ is a vector of six lagged observed variables (from x_{t-5} to x_t). The within-person component is defined as:

$$\mathbf{x}_{wit}^{(6)} = \mathbf{L} \mathbf{g}_{it} + \boldsymbol{\epsilon}_{it} \quad (10)$$

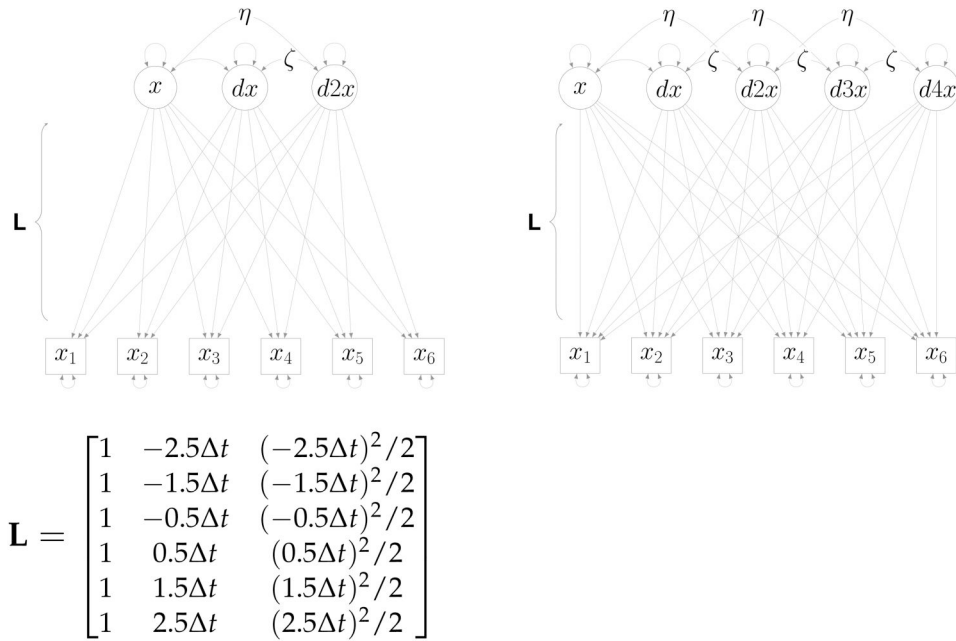


Figure 3. Standard second-order vs. fourth-order LDSEM.

Note. The first diagram represents the standard second-order LDSEM model. The second diagram represents the fourth-order LDSEM model. $\mathbf{L}$ represents the specialized factor loadings that constrain the relationships between the observed variables (e.g., x_1 to x_6) and the latent derivatives. x refers to the smoothed level of the variable x . dx , $d2x$, $d3x$, and $d4x$ refer to the first to the fourth derivative of x . We show an example of $\mathbf{L}$ with $d=6$, $\Delta t=1$, and the highest order = 2. This figure does not encapsulate the multilevel component.

Consistent with the formulation of multiple regression models within the structural equation modeling framework in which all independent variables are allowed to freely covary, x and dx are also allowed to freely covary with each other.

where $\mathbf{L}$ is a fixed factor loading matrix¹ shown in Figure 3 and $\mathbf{g}_{it}$ is a vector of unobserved latent derivative scores, $[x_{it}, \dot{x}_{it}, \ddot{x}_{it}]'$. The structural equations for $\mathbf{g}_{it}$ are represented as:

$$\mathbf{g}_{it} = \mathbf{A}_i \mathbf{g}_{it} + \mathbf{e}_{it} \quad (11)$$

$$\mathbf{A}_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \eta_i & \zeta_i & 0 \end{bmatrix}$$

where η_i and ζ_i are person-specific parameters potentially influencing daily fluctuations in positive emotions as described in Equation (1).

For the between-person component, we have:

$$\begin{bmatrix} \eta_i \\ \zeta_i \end{bmatrix} = \begin{bmatrix} \eta_0 \\ \zeta_0 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} Cov_{Bi} + \zeta_{Bi} \quad (12)$$

where the intercepts (η_0, ζ_0) are the predicted values for the frequency and damping parameters, respectively when the covariate (Cov_{Bi}) is zero, and the regression coefficients (β_1, β_2) indicate changes in these parameter values with each unit of increase in the person-specific covariate, Cov_{Bi} . Note that in our

proposed model, $\mathbf{x}_{Bi}^{(6)}$ in Equation (9) is simply a vector of zeros because we assumed the baseline of all individuals' processes is fixed at zero, thus rendering their intercepts 0 in the corresponding time delay embedded data. However, this can be further modeled as described in a regular MSEM. For a more detailed exploration of MSEM, please refer to Depaoli and Clifton (2015) and Hox (2013).

Leveraging higher-derivative information

One determinant of the performance of the M-LDSEM is whether the time-delay embedding dimension is well selected. In addition to that, specifying the extraction of higher-order derivatives (i.e., up to the fourth order) in an M-LDSEM, as proposed by Boker et al. (2020), can significantly improve the performance of the approach (Chow, 2019), even if the dimension of time-delay embedding is less than optimal.

The 4th-order LDSEM entails the use of higher-order derivative information in the Taylor series approximation of the observed measurements, which is a second-order differential equation including the 3rd or 4th-order derivatives as a dependent variable in the model as shown in the second diagram in

¹ $\mathbf{L}$ can in principle be person- and time-specific when the time intervals between successive occasions are different across individuals and over time.

Figure 3. This is feasible by using specialized factor loading (L) described above and some invariance constraints imposed across derivative orders such that the frequency parameters (η) and damping parameters (ζ) in each equation are constrained to be equal, respectively. More details about the 4th-order LDSEM can be found in Boker et al. (2020).

Simulation study

The main purposes of the current simulation study are to compare the relative performance of Bayesian and robust Frequentist estimation approaches, and clarify the benefits of including higher-order derivative information. Specifically, we were interested in fitting differential equation models with finite-length data, time-varying covariates, and coupling effects. With these goals in mind, three design factors (i.e., derivative order, estimation approach, and data configuration) were varied within our simulation study.

Simulation conditions

Modeling approach

As modeling approaches, 2 derivative orders $\times$ 3 estimation approaches = 6 conditions were considered. First, two derivative order specifications were used as described in Figures 3: (1) second-order and (2) fourth-order M-LDSEM. Then, three estimation approaches (i.e., two frequentist and one Bayesian approaches) were considered: (1) single-level robust estimator, (2) two-level robust estimator, and (3) two-level Bayesian estimator.

For the single-level robust estimator, the estimator MLR and TYPE=COMPLEX commands in Mplus were used to obtain maximum likelihood estimates with standard errors robust to non-normality and non-independence (Asparouhov, 2005). This single-level approach does not account for the random effects in the data generation model but offers corrections to the standard error estimates under such model misspecification through the Hubert-White sandwich estimator (Freedman, 2006).

The other two-level approaches allow for the estimation of random effects, as consistent with our data generation model. For the two-level robust approach, the estimator MLR and TYPE=TWOLEVEL commands in Mplus were used to compute maximum likelihood robust standard errors (Asparouhov & Muthen, 2006) with Monte Carlo integration to handle estimation of the random effects. The two-level Bayesian estimator condition used the Markov Chain

Monte Carlo (MCMC) algorithm based on the Gibbs sampler (Gelman et al., 1995), with default settings for Mplus (see Asparouhov & Muthen, 2010). Specifically, conjugate priors were used². Convergence was determined based on the potential scale reduction (PSR) values calculated from two MCMC chains, with PSR values ≤ 1.1 across all parameters used as a cutoff for convergence. The first half of each chain was designated as burn-in iterations and not utilized for estimation purposes. The point estimates were then derived from the median of the MCMC samples obtained from the second half of each chain.

Sample size configuration

The sample size configuration was decided based on our motivating empirical data ($T=22$, $N=140$) and our key interest in elucidating the effects of the length of the series, especially in scenarios involving short, multi-subject time series. Three conditions of the length of the series were considered. $T=100$ corresponds to typical time series lengths seen in other simulation studies involving time-intensive dynamical systems models and data (Boker et al., 2020; Chow, 2019). $T=20$ and 10 were picked to evaluate how well the M-LDSEM approach worked under sample size configurations that mirrored those seen in common EMA designs (Konjarski et al., 2018; Rodríguez-Blanco et al., 2018). The number of subjects was set to $n=150$ taking into account characteristics of our motivating empirical data, and common EMA sample sizes as reported in recent review papers of EMA studies (mean number of subjects = 146.3; median = 87; Konjarski et al., 2018; Liu et al., 2019; Yang et al., 2019). Additionally, a condition with $n=300$ was additionally tested for $T=10$ to examine the effect of doubling the number of subjects. Consequently, four sample size configurations were considered: (1) $n=150$ & $T=100$, (2) $n=150$ & $T=20$, (3) $n=150$ & $T=10$, and (4) $n=300$ & $T=10$.

Data generation and implementation of M-LDSEM

We have 2 derivative orders $\times$ 3 modeling approaches $\times$ 4 sample size configurations = 24 simulation conditions. In each condition, 500 Monte Carlo replications were simulated. Specifically, data following the coupled damped oscillator model in Equation (13) were simulated in R (R Core Team, 2022) using the

²The default prior for intercepts, loadings, and slopes for continuous variables is $N(0, 10^{10})$. The default prior for variance-covariance matrices for continuous variables is $IW(0, -p-1)$, where p is the number of observed continuous variables in the model.

lsoda function in the R package, deSolve (Soetaert et al., 2010).

$$\begin{aligned}\ddot{x}_{it} &= \eta_1 x_{it} + \zeta_1 \dot{x}_{it} + \gamma_1 y_{it} + \beta_{1i} XTVcov_{it} + e_{x,it} \\ \beta_{1i} &= \beta_{x0} + \beta_{x1} Cov_i + u_{xi} \\ \ddot{y}_{it} &= \eta_2 y_{it} + \zeta_2 \dot{y}_{it} + \gamma_2 x_{it} + \beta_{2i} YTVcov_{it} + e_{y,it} \\ \beta_{2i} &= \beta_{y0} + \beta_{y1} Cov_i + u_{yi}\end{aligned}\quad (13)$$

where x_{it} and y_{it} are values of variables of interest (e.g., affect of patients and spouses) at time t for dyad i ; $XTVcov_{it}$ and $YTVcov_{it}$ are time-varying covariates (e.g., daily instrumental support received/provided); Cov_i represents a time-invariant covariate; and u_{xi} & $u_{yi} \sim N(0, \tau^2)^3$ and $e_{x,it}$ & $e_{y,it} \sim N(0, \sigma^2)$. For all simulation conditions, the model parameters were specified as follows: $\eta_1 = \eta_2 = -0.5$, $\zeta_1 = \zeta_2 = -0.04$, $\gamma_1 = \gamma_2 = -0.1$, $\beta_{x0} = \beta_{y0} = 0.3$, $\beta_{x1} = \beta_{y1} = 0.1$, $\tau^2 = 0.04$, and $\sigma^2 = 1$. The simulation parameters were selected to reflect the estimated values from previous studies using similar affect variables (Chow et al., 2005; Hilpert et al., 2020; Steele & Ferrer, 2011). Specifically, previous studies reported that η was estimated to be between -0.4 to -0.9 , while ζ was either insignificant or very small, ranging from -0.01 to -0.04 . We assumed that a single manifest indicator was used, and the baseline was fixed at zero for everyone.

The time-invariant covariate was simulated from a uniform distribution over $[-3, 3]$. Two time-varying covariates were simulated separately following the stochastic Ornstein–Uhlenbeck (OU) process with zero home-base (Oravecz et al., 2011) as: $dx_{it} = -\theta x_{it} dt + \phi dW_{it}$, where x_{it} is the value of the time-varying covariate of person i at time t , W_t is the standard Wiener process, $\theta = 0.03$, and $\phi \sim N(0, 0.1)$. Although detrending is a commonly adapted data preparation step (e.g., Chow et al., 2005; Hilpert et al., 2020), we did not detrend in this case, because the simulated time-varying covariate following the OU process shows a linear time trend. Detrending in this case would remove key information related to the time-varying covariate and would not be appropriate for the purposes of the current study. We verified, *via* a targeted Monte Carlo simulation, that removal of a linear time trend would not lead to biases in the estimation results when the time-varying covariate

assumed other alternative forms of change, such as a sine function (see Appendix A for details).

Following Step 1 of the M-LDSEM approach, a time-delay embedded matrix was created for each Monte Carlo replication. Consonant with the settings adopted in the empirical example, the dimension of the time delay embedding, d , was set to six. Latent factors were also specified for time-varying covariates with $d = 6$. Please note that d does not necessarily need to always be six. Detailed rationales and guidelines for selecting this particular value of d are provided in Empirical Example.

Performance measures

Analyses were conducted using Mplus 8.7 (Muthen & Muthen, 2017) and an R package, Mplusautomation (Hallquist & Wiley, 2018). Performance measures were calculated for the 500 replications in each condition. To quantify the performance of the point estimates, the mean point estimate, relative bias, and root mean squared error (RMSE) were used. To quantify the variability of the estimation, the standard deviation of the parameter estimates in all Monte Carlo simulations (MCSD; Monte Carlo Standard Deviation) was used. As a measure of the relative performance of the SE estimates, the relative deviance of the standard error (RDSE) was calculated based on the difference between the SE estimate and the empirical standard deviation of each parameter (the “true” SE; MCSD). The estimated SE was calculated as the average of the SE estimates from the model results ($\widehat{aSE}$). Power was computed by tallying the proportion of all trials in which the 95 % CIs did not include zero. Details of these performance measures are included as footnotes in Tables 2 and 3.

Simulation results

Are there benefits to including higher-order derivatives?

Models using Single-level robust and the two-level Bayesian approaches converged 100% across all Monte Carlo replications regardless of the derivative order or sample size configuration. Under the two-level robust approach, convergence rates varied between 49 to 86.2%, as dependent on the highest derivative order of the M-LDSEM and sample configuration. As shown in Table 1, the length of the time series was a critical factor that influenced the convergence rate; even though doubling the sample size from $n = 150$ to 300 helped increase the convergence rates, data with

³Note that we only allowed random variance for the time-varying covariates in our simulation for simplicity as this study is an initial comparison of the Bayesian and frequentist approaches using M-LDSEM. However, it is possible to specify more complex random effect structures in practice, such as allowing η s, ζ s, and γ s to vary across dyads. This was demonstrated in the Empirical Example section where we allowed inter-dyad differences in coupling parameters.

$T=10$ resulted in convergence rates that were around 50%. Despite the complexity of the 4th order model, inclusion of higher derivatives for the same n and T configurations led to higher convergence rates than only using the same order of derivatives (second derivatives) as the data generation model.

Other key simulation results comparing the 2nd to 4th derivative order are presented in Table 2.

Table 1. Convergence rates of the two-level robust modeling approach.

Conditions Order	n	T	Convergence rate
2	150	100	85.6%
2	150	20	79.6%
2	150	10	49.0%
2	300	10	64.6%
4	150	100	86.2%
4	150	20	85.6%
4	150	10	53.6%
4	300	10	72.6%

Note. Convergence rates were calculated as the percentage of the 500 replications that did not encounter convergence errors. Convergence errors resulted from either the failure to find an optimized solution or the optimizer exceeding the maximum allotted iterations.

Performance measures for parameters of the same type (e.g., coupling parameters for patients and spouses) were aggregated to ease presentation. In terms of point estimates, the 4th-order models led to smaller biases than the 2nd-order models for all parameters except for the damping parameter (ζ demonstrated a minimal relative bias, -0.03 , when the highest derivative order was 2). In contrast to the point estimates, the 2nd-order models resulted in smaller MCSD than the 4th-order models for all parameters but ζ . This suggested that inclusion of higher derivatives helped improve convergence rates and point estimates for most parameters, even though some overestimation in SEs was observed under the 4th compared to 2nd-order models particularly under the multilevel robust frequentist approach.

How does the Bayesian approach perform relative to the frequentist approaches?

When the three estimation approaches were compared, the two-level Bayesian estimator resulted in the most accurate estimates with smaller relative biases

Table 2. Summary statistics of parameter estimates across model derivative orders and modeling approaches across 500 Monte Carlo replications when $n = 150$, $T = 100$.

	Order	Estimator	θ	Mean $\hat{\theta}$	rBias	RMSE	MCSD	$\widehat{aSE}$	RDSE
η	2	2-Bayes	-0.50	-0.38	-0.23	0.116	0.0020	0.0020	0.010
η	2	2-Robust	-0.50	-0.34	-0.32	0.160	0.0151	0.0168	0.110
η	2	1-Robust	-0.50	-0.30	-0.40	0.200	0.0125	0.0123	-0.013
η	4	2-Bayes	-0.50	-0.50	0.00	0.005	0.0047	0.0051	0.095
η	4	2-Robust	-0.50	-0.43	-0.15	0.076	0.0223	0.0341	0.531
η	4	1-Robust	-0.50	-0.38	-0.25	0.126	0.0166	0.0165	-0.011
ζ	2	2-Bayes	-0.04	-0.04	-0.03	0.003	0.0028	0.0043	0.503
ζ	2	2-Robust	-0.04	-0.03	-0.20	0.009	0.0034	0.0060	0.760
ζ	2	1-Robust	-0.04	-0.03	-0.23	0.010	0.0030	0.0031	0.031
ζ	4	2-Bayes	-0.04	-0.02	-0.41	0.017	0.0023	0.0036	0.574
ζ	4	2-Robust	-0.04	-0.02	-0.50	0.020	0.0025	0.0032	0.258
ζ	4	1-Robust	-0.04	-0.02	-0.51	0.021	0.0024	0.0024	0.022
γ	2	2-Bayes	-0.10	-0.07	-0.33	0.033	0.0019	0.0017	-0.128
γ	2	2-Robust	-0.10	-0.06	-0.36	0.036	0.0048	0.0067	0.409
γ	2	1-Robust	-0.10	-0.06	-0.42	0.042	0.0055	0.0054	-0.013
γ	4	2-Bayes	-0.10	-0.08	-0.17	0.017	0.0027	0.0022	-0.188
γ	4	2-Robust	-0.10	-0.08	-0.22	0.023	0.0060	0.0109	0.806
γ	4	1-Robust	-0.10	-0.07	-0.28	0.028	0.0066	0.0065	-0.011
β_0	2	2-Bayes	0.30	0.23	-0.22	0.066	0.0133	0.0135	0.014
β_0	2	2-Robust	0.30	0.22	-0.27	0.084	0.0194	0.0179	-0.076
β_0	2	1-Robust	0.30	0.19	-0.36	0.110	0.0158	0.0150	-0.053
β_0	4	2-Bayes	0.30	0.31	0.03	0.019	0.0173	0.0176	0.019
β_0	4	2-Robust	0.30	0.28	-0.08	0.036	0.0258	0.0313	0.212
β_0	4	1-Robust	0.30	0.24	-0.20	0.064	0.0205	0.0194	-0.053
β_1	2	2-Bayes	0.10	0.08	-0.21	0.023	0.0075	0.0078	0.032
β_1	2	2-Robust	0.10	0.07	-0.27	0.029	0.0093	0.0117	0.260
β_1	2	1-Robust	0.10	0.06	-0.36	0.037	0.0080	0.0079	-0.012
β_1	4	2-Bayes	0.10	0.10	0.03	0.010	0.0099	0.0104	0.049
β_1	4	2-Robust	0.10	0.09	-0.08	0.014	0.0122	0.0142	0.163
β_1	4	1-Robust	0.10	0.08	-0.20	0.022	0.0101	0.0100	-0.010
τ^2	2	2-Bayes	0.04	0.03	-0.34	0.014	0.0030	0.0032	0.089
τ^2	2	2-Robust	0.04	0.01	-0.66	0.027	0.0046	0.0031	-0.325
τ^2	4	2-Bayes	0.04	0.04	0.14	0.007	0.0051	0.0056	0.096
τ^2	4	2-Robust	0.04	0.02	-0.45	0.019	0.0074	0.0051	-0.309

Note. 1-Robust: single-level robust estimator; 2-Robust: two-level robust estimator; 2-Bayes: two-level Bayesian estimator; θ = true value of a parameter; Mean $\hat{\theta} = \frac{1}{H} \sum_{h=1}^H \hat{\theta}_h$, where $\hat{\theta}_h$ = estimate of θ from the h th Monte Carlo (MC) runs; rBias = relative bias = $\frac{1}{H} \sum_{h=1}^H (\hat{\theta}_h - \text{true}\theta) / \text{true}\theta$; RMSE = $\sqrt{\frac{1}{H} \sum_{h=1}^H (\hat{\theta}_h - \text{true}\theta)^2}$; MCSD = standard deviation of $\hat{\theta}$ across MC runs, aSE = average standard error estimate across MC runs, $RDSE$ = average relative deviance of $SE = (aSE - MCSD) / MCSD$.

Table 3. Summary statistics of the parameter estimates with two-level Bayesian estimation and 4th-order model across 500 Monte Carlo replications.

	T	n	θ	Mean $\hat{\theta}$	rBias	RMSE	MCSD	$\widehat{aSE}$	RDSE	Power	Cvr
η	100	150	-0.50	-0.500	0.000	0.005	0.0047	0.0051	0.0946	100	95.8
η	20	150	-0.50	-0.505	0.009	0.009	0.0082	0.0073	-0.1154	100	84.2
η	10	150	-0.50	-0.508	0.016	0.018	0.0160	0.0105	-0.3403	100	74.1
η	10	300	-0.50	-0.505	0.011	0.012	0.0111	0.0072	-0.3534	100	72.1
ζ	100	150	-0.04	-0.024	-0.411	0.017	0.0023	0.0036	0.5744	100	0
ζ	20	150	-0.04	-0.028	-0.309	0.014	0.0073	0.0093	0.2706	90	77.4
ζ	10	150	-0.04	-0.023	-0.426	0.024	0.0165	0.0128	-0.2254	44	69.5
ζ	10	300	-0.04	-0.023	-0.415	0.021	0.0123	0.0098	-0.2022	65	56.9
γ	100	150	-0.10	-0.083	-0.168	0.017	0.0027	0.0022	-0.1880	100	0
γ	20	150	-0.10	-0.079	-0.206	0.021	0.0048	0.0046	-0.0383	100	0.8
γ	10	150	-0.10	-0.081	-0.194	0.021	0.0092	0.0078	-0.1458	100	30.9
γ	10	300	-0.10	-0.083	-0.174	0.018	0.0062	0.0047	-0.2374	100	7.4
β_0	100	150	0.30	0.308	0.028	0.019	0.0173	0.0176	0.0188	100	93.4
β_0	20	150	0.30	0.313	0.044	0.025	0.0218	0.0218	0.0012	100	91
β_0	10	150	0.30	0.311	0.037	0.029	0.0269	0.0258	-0.0424	100	92.8
β_0	10	300	0.30	0.309	0.030	0.021	0.0188	0.0179	-0.0455	100	89.8
β_1	100	150	0.10	0.103	0.029	0.010	0.0099	0.0104	0.0492	100	95.6
β_1	20	150	0.10	0.104	0.038	0.013	0.0123	0.0126	0.0247	100	94.8
β_1	10	150	0.10	0.103	0.029	0.015	0.0144	0.0146	0.0111	100	95.2
β_1	10	300	0.10	0.103	0.028	0.010	0.0099	0.0101	0.0213	100	95.6
τ^2	100	150	0.04	0.045	0.137	0.007	0.0051	0.0056	0.0957	100	87.8
τ^2	20	150	0.04	0.057	0.419	0.019	0.0089	0.0083	-0.0661	100	49.8
τ^2	10	150	0.04	0.059	0.468	0.022	0.0121	0.0113	-0.0652	100	69.9
τ^2	10	300	0.04	0.057	0.419	0.019	0.0082	0.0076	-0.0831	100	37.3

Note. θ = true value of a parameter; Mean $\hat{\theta} = \frac{1}{H} \sum_{h=1}^H \hat{\theta}_h$, where $\hat{\theta}_h$ = estimate of θ from the h th Monte Carlo (MC) runs; rBias = relative bias = $\frac{1}{H} \sum_{h=1}^H (\hat{\theta}_h - \text{true}\theta) / \text{true}\theta$; RMSE = $\frac{1}{H} \sum_{h=1}^H (\hat{\theta}_h - \text{true}\theta)$; MCSD = standard deviation of $\hat{\theta}$ across MC runs; $\widehat{aSE}$ = average standard error estimate across MC runs; RDSE = average relative deviance of $\widehat{SE} = (aSE - SE)/SE$; Power = 1 - the proportion of 95% credible intervals that contain 0 across MC runs; Cvr = Coverage = proportion of 95% credible intervals that contain θ across MC runs.

and RMSEs. The notably higher relative bias for ζ was accentuated in part by the division of a small average bias (0.02) by the near-zero true value of ζ (0.04). In addition, the relative bias for ζ decreased considerably from -0.41 for the 4th-order model to -0.03 for the 2nd-order model. This suggests that, for ζ , using 4th-order information in the Bayesian model extracted too much nuanced fluctuations in the data that led to underestimation of the amount of true damping. Still, given the relatively superior performance of the 4th-order approach for all other parameters except for ζ , we would still recommend use of the 4th over the 2nd-order derivative information. The Bayesian estimator also yielded the best efficiency (smaller magnitudes of $\widehat{aSE}$ and MCSD were regarded as higher efficiency), for all but the damping parameters, as well as accurate quantification of the uncertainty around these estimates (based on smaller RDSE). It is noteworthy that inclusion of 4th-order derivatives in conjunction with the Bayesian estimator led to notable improvements in estimation quality in general, especially for coupling parameters when compared to previous simulation results utilizing the 2nd-order LDSEM method (Hu et al., 2014).

In contrast to the Bayesian estimator, other frequentist approaches were characterized by relatively high biases, especially in the parameters η , γ , and the random effect variances (τ^2). As expected, the single-level frequentist approach with sandwich-type SE

estimator was associated with the greatest biases compared to other approaches that did incorporate the correctly specified random effects structure. However, this approach yielded reasonably accurate SE estimates, with smaller RDSEs compared to the multilevel robust frequentist approach, which was frequently characterized by overestimation in SE estimates, and even the Bayesian estimator for the parameters ζ and γ . The power estimates were close to 100% in all conditions and are thus omitted from Table 2.

The efficacy of the DIC as Bayesian model selection criteria in M-LDSEM

Although the use of DIC with a time-delayed embedding matrix is not well-established, the Deviance Information Criterion (DIC) is a well-known model fit index in the Bayesian framework, where a smaller DIC indicates a better fit. Therefore, we explored the use of DIC in M-LDSEM and found that the DIC consistently favored the 4th-order model over the standard 2nd-order model across all sample size configurations.

To further examine the usefulness of DIC in M-LDSEM, we conducted two additional simulations with 500 replications with $T = 100$ and $n = 150$. Our findings suggested that using DIC as a measure of model fit in M-LDSEM may have limitations. Firstly, we fitted a “reduced” model to our simulation data by omitting the estimation of the effect of a time-varying

covariate. Ideally, if DIC were effective, the true “full” model, which we used in our original simulation, should have a smaller DIC. However, only 19.2% of cases correctly selected the true model as the preferred model. Secondly, when an “over-parameterized” model that included a time-varying covariate’s effect was fitted to data generated without any time-varying covariate, the correct (less complex) model was preferred by the DIC in only 68.2% of the replications. Given our findings and the weak theoretical justification for DIC (Spiegelhalter et al., 2014), caution is necessary when using DIC as a model selection criterion in M-LDSEM.

Can model parameters be recovered with short time series?

In this section, we focus on addressing the relative performance of the “best-performing” approach, namely, the Bayesian estimation approach under 4th-order derivatives, under three time series length conditions and n of 150 and 300. Table 3 shows the summary statistics of the results. Overall, even though greater biases in all parameters were observed under small T , the relative biases and RMSEs of all parameters were still relatively low compared to comparable results observed earlier for $T=100$ and $n=150$ in Table 2 under the frequentist approaches. For inferences involving the covariate-related regression coefficients (β_0 and β_1), relative biases, RMSEs, and RDSEs were particularly low.

For η and ζ , doubling T from 10 to 20 led to a greater reduction in RMSEs than doubling n from 150 to 300. Among the parameters associated with the time-varying covariate effects (β_0 and β_1), γ and τ^2 , slightly greater decreases in RMSEs were observed in doubling n than in doubling T . Recovery of the damping parameters was particularly challenging; the results for ζ were biased for all time lengths. It is not surprising considering that the damping parameter was less biased under the conventional 2nd-order LDSEM model. More importantly, unlike other parameters, ζ only reached 44% of power when $T=10$. This was in stark contrast to all other parameters, whose power estimates were close to or at 100% even with $T=10$.

Despite the decent coverage rates for most variables, the coverage rates for ζ and γ remained low even under the most ideal conditions. We observed that while biases decreased for other parameters as the number of time points increased, the biases for ζ and γ remained constant. In contrast, the standard error for these parameters decreased as expected. Consequently, the underestimation of the standard

error combined with the biases led to even worse coverage rates with an increase in the number of time points.

Summary of simulation results

Overall, the present study validated the relative gains in estimation quality when higher-order derivative information was included in M-LDSEM. Also, the Bayesian estimator outperformed other frequentist approaches in general. Combining the 4th-order model and the Bayesian estimator yielded relatively good estimation quality even with short time series such as $T=10$ or 20. Understandably, greater T still played a critical role in improving estimation properties, but the relatively high power estimates provided some reassurance of the feasibility of making inferences on intra- and inter-personal dynamics and corresponding predictors in the context of relatively small T . Further, considering the low model convergence rates when the two-level robust approach was used (see Table 1), the practical utility of the Bayesian estimator was especially pronounced when T is short.

Empirical example

In our simulation study, we verified that use of the Bayesian estimator combined with inclusion of higher derivative information performed relatively well even with short, multiple-subject time series data. To show the practical utility of the proposed approach, we provide an illustrative example using the 22-day daily diary data of 140 couples of OA patients and their spouses (Martire et al., 2013). Of the 22 days, the median completion rate was 17 days. Positive affect of the couples was measured three times a day, and these measurements were aggregated into daily positive affect scores for each dyad member. Spousal instrumental support was reported by both patients and spouses at the end of the day (e.g., whether the I/spouse tried to get the patient to rest when they seemed to be in pain).

Empirical analysis

As described in section Motivating Example, we sought to investigate how spousal instrumental support was related to patients’ and spouses’ positive affect. Caring for chronically ill partners is a complex phenomenon for both providers and recipients; the consequences of support are not always positive and differ across couples (Beach et al., 2000; Newton-John, 2013; Reinhardt et al., 2006). Although previous

studies have used primarily day-to-day covariance as an indication of the average daily association between instrumental support and couple's affect (Gremore et al., 2011; Marini et al., 2021), few studies have provided adequate consideration of how each dyad member's emotions evolve over time and the degree to which patients' and spouses' patterns of change covary with one another (Sbarra & Hazan, 2008; Sels et al., 2018). Thus, following insights from the simulation study, we fitted a bivariate oscillator model in Equation (4). Utilizing this model also allows researchers and practitioners to disentangle the patient $\rightarrow$ spouse and spouse $\rightarrow$ patient influences while elucidating the role of instrumental support.

In Equation (4), the instrumental support that patients or spouses reported was used as a time-varying covariate in the patient's or spouse's equation, respectively. To limit modeling complexity associated with the number of random effects, random effects were included only for the coupling parameters (γ_{pi} and γ_{si}) as well as the regression parameters for instrumental support (β_{pi} and β_{si}), given our key interest in understanding the influences between patients and their caregivers. Further, we tested whether the typical level of instrumental support in each couple indirectly influences the couple's positive affect as moderators of the coupling effect and the effect of daily support. All time-varying variables (i.e., $Patient_{it}$, $Spouse_{it}$, $PatientSupport_{it}$, and $SpouseSupport_{it}$) were centered within subjects.⁴ The average instrumental support scores were grand mean centered and standardized across individuals. Since we did not expect any damping over the 22 days and the damping parameters were not different from zero in an initial model fitting, they were fixed at zero and omitted in Equation (4).

The model was fitted using a six-dimensional time-delayed embedded matrix consisting of data at time $t_0, t_1, \dots, t_5$. This imposed a constraint that at least 5 time points from each dyad are needed to construct the input data matrix. To allow for at least 3 replicated rows of time-delay embedded data from each dyad for inference of interindividual differences, we retained data only from dyads with at least 8 days of measurements, yielding a total of $N = 121$ dyads for model fitting purposes. As illustrated, available input data (i.e., the number of rows in the time-delay embedded matrix) decreases as d increases. When

deciding on an embedding dimension (d), this may not be critical for time series with ample time points, but for short time series with many missing data, it could be a crucial factor. Therefore, we recommend considering the remaining sample size depending on the selected d .

We also recommend considering the characteristics of the dynamics of interest based on previous studies. While a higher d will be more robust to noisy data, for data with relatively salient signals and low noise, a smaller d would be sufficient. For instance, daily affective assessment may not have high noise as typical physiological signal processes have. Thereby, studies that examined daily emotion regulation have typically used d of 4, 5, or 6 (Chow et al., 2005; Hilpert et al., 2020; Steele & Ferrer, 2011). Additionally, Boker et al. (2020)'s experiment with d values suggests that $d = 6$ is appropriate for short time series ($T = 50$), given processes with $\eta = -0.5$, $\zeta = -0.1$, which is close to the expected range of the frequency of affect dynamics.

While it has been demonstrated that the 4th order LDSEM is less sensitive to the selection of embedding dimension (Boker et al., 2020), we acknowledge that more complex models, such as the coupled oscillator model with multilevel structure, may still be sensitive to the choice of embedding dimension and require further research.

Before proceeding to the empirical result, we acknowledge that the empirical model has more complexity than the main simulation model in the sense that it has random effects for their coupling parameters. To closely resemble the complexity of the empirical model, we additionally conducted a simulation study that incorporated individual variability in the coupling parameter and used zero damping parameter, which is presented in Appendix B. Consistent with our main simulations, we observed an underestimation of the coupling parameter and its related covariate effect. As anticipated, introducing individual differences in the coupling parameters made the estimation slightly more challenging. Thus, more caution is needed to interpret the coupling parameters when they are allowed to differ across participants. However, the directionality of the effects remained accurate, and the power was consistently at 100% for all parameters. Despite the integration of random effects and the increased model complexity, the recovery of the parameters related to time-varying covariates remained commendable. Hence, while our method may not be flawless, it still provides valuable

⁴Please note that centering was unnecessary for our simulation since each time series was generated to fluctuate around zero. However, if the time-varying covariate is not already centered at zero, it can be centered within individual, as we did in our empirical example.

Table 4. Parameter estimates from the empirical model.

Parameter	Description	Est.	Post. SD	95% CI	
				LL	UL
η_p	Frequency	-0.78	0.03	-0.83	-0.73
$\gamma_{pi} = \gamma_{p0} + \gamma_{p1}\text{PatientSupport}_i$	Coupling				
γ_{p0}		0.13	0.06	0.01	0.26
γ_{p1}		-0.11	0.11	-0.33	0.11
$\beta_{pi} = \beta_{p0} + \beta_{p1}\text{PatientSupport}_i$	Support				
β_{p0}		0.05	0.08	-0.11	0.21
β_{p1}		0.13	0.19	-0.26	0.51
η_s	Frequency	-0.87	0.03	-0.93	-0.80
$\gamma_{si} = \gamma_{s0} + \gamma_{s1}\text{PatientSupport}_i$	Coupling				
γ_{s0}		0.09	0.05	0.01	0.19
γ_{s1}		0.11	0.11	-0.10	0.32
$\beta_{si} = \beta_{s0} + \beta_{s1}\text{PatientSupport}_i$	Support				
β_{s0}		0.19	0.08	0.03	0.35
β_{s1}		-0.51	0.23	-0.95	-0.03
Random Effect					
$\text{var}(u_{\gamma, pi})$		0.19	0.05	0.12	0.32
$\text{var}(u_{\gamma, si})$		0.14	0.03	0.09	0.22
$\text{cov}(u_{\gamma, pi}, u_{\gamma, si})$		0.06	0.02	0.02	0.11
$\text{var}(u_{\beta, pi})$		0.45	0.09	0.31	0.67
$\text{var}(u_{\beta, si})$		0.40	0.10	0.25	0.66
$\text{cov}(u_{\beta, pi}, u_{\beta, si})$		0.11	0.07	-0.01	0.26
σ_p^2		0.25	0.01	0.24	0.26
σ_s^2		0.24	0.01	0.23	0.25

Note. Post. SD = Standard deviation of the posterior distribution; LL = lower limit; UL = upper limit.

insights, particularly when considering its computational efficiency.

Empirical results

The results are presented in Table 4. Consonant with findings in affect literature (Bonanno, 2001; Chow et al., 2005), the positive affect of couples was found to be oscillatory in nature, as indicated by the negative frequency parameters (η s). In line with interpersonal emotion regulation theories (Elfenbein, 2014; Niven, 2017), the coupling effects (γ s) were positive. The significant coupling effects in both the directions of *Patient* $\rightarrow$ *Spouse* and *Spouse* $\rightarrow$ *Patient* indicated that on days with typical amounts of instrumental support (see the middle plot of (a) in Figure 4), the patients and spouses reported reciprocal influences on each other. In other words, when one dyad member reported higher-than-usual levels of positive affect, such positive deviations in affect accelerated and elevated the partner's positive affect. Significant between-dyad differences were found in the magnitudes of these coupling effects; however, the typical level of instrumental support did not explain these differences.

The effect of daily instrumental support was only significant for the spouses, even though significant between-couple differences were found in the magnitudes of this covariate on patients as well as spouses. Since the daily support variable was within-person mean-centered, a high (> 0) value of support variable represents a day on which the spousal instrumental

support was high relative to the spouse's typical level of support (e.g., average support over the 22 days). Thus, our result suggested that, on the days when the spouses reported providing a higher-than-usual level of instrumental support (i.e., +1SD), greater accelerations in their positive affect were evidenced, resulting in more extreme, positive upward shifts in their positive affect over time. See Figure 4(a) for the simulated trajectories. This outcome is in line with an earlier finding that suggests supporting their loved ones may be beneficial to their own well-being (Poulin et al., 2010). Our result further clarified that contrary to the direct effect characterized by provision of instrumental support on spouses, the patients' positive affect was not significantly associated with variations in daily support received. Rather, patients' positive affect showed greater fluctuations (in both the positive and negative directions), presumably due to the positive coupling effect on their spouses.

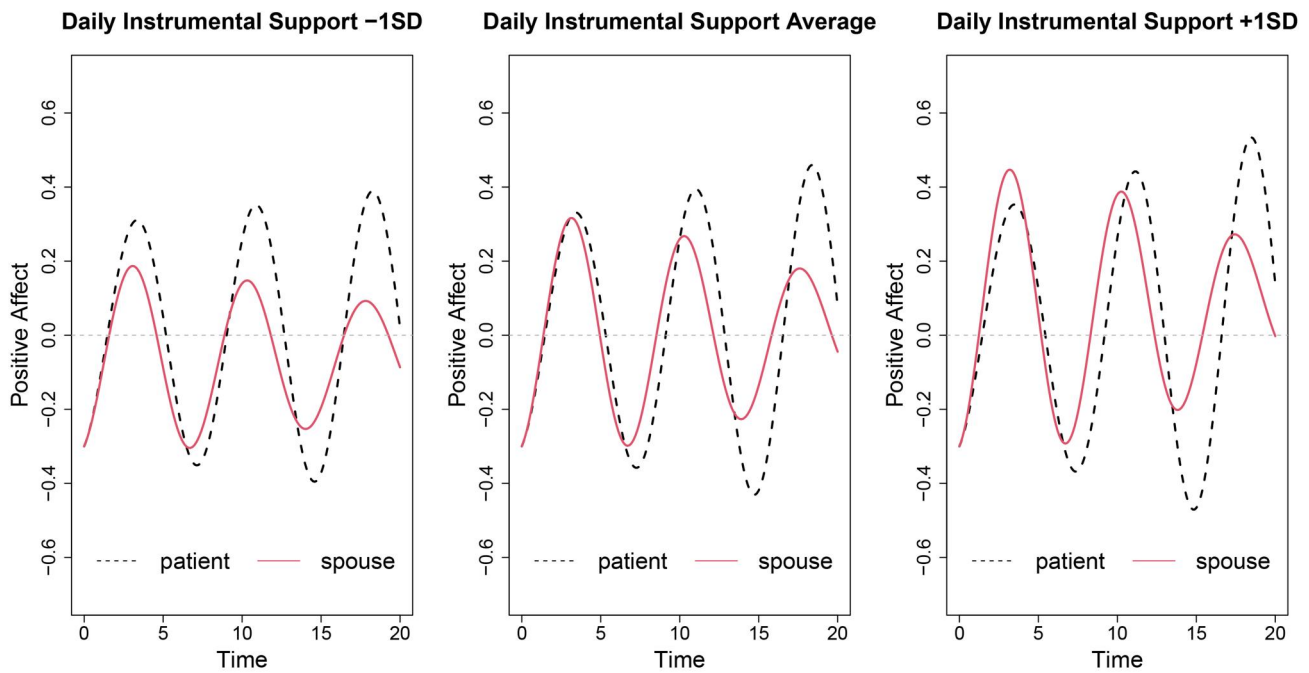
To further evaluate the robustness of our significant findings, we conducted additional simulations to examine Type I error rates when using the proposed approach. This time, we fixed either the effect of daily instrumental support or the moderation effect of the average levels of support for the daily support's effect at zero, respectively. The results showed a Type I error rate of 3.6% for the fixed effect of daily support and 5.6% for the fixed moderation effect. These results suggested that the proposed approach yielded Type I error rates that were close to the nominal rate of 5%.

Finally, the effect of daily support for spouses varied depending on their average support levels ($\hat{\beta}_{s1} = -0.51$). Figure 4(b) shows that, for spouses who typically provide high instrumental support (+1SD), higher-than-average provision of daily support did accelerate their positive emotions, but the effect was weaker compared to those whose typical support level was low (-1SD). It is in line with our hypothesis that for those who are used to providing abundant support, the effect of daily ups/downs of the support provision would be reduced. This may shed light on past inconsistent results on the impact of daily instrumental support by explaining the inter-couple differences.

As a rudimentary measure of effect sizes for the proposed M-LDSEM, we report the averaged level-1 R^2 across clusters (individuals) provided by MPlus⁵. In our study, we used a 6-dimensional embedded

⁵Since R^2 at the within-person level (level-1) is not a constant value but rather, is person-specific if there are random slopes and effects of person-specific predictors in the model, Mplus provides R^2 that is within-level averaged over clusters (Schuurman et al., 2016). Furthermore, Mplus provides an R^2 for each dependent variable in the model, without breaking it down by covariate.

(a) Trajectories of positive affect at low ($-1SD$), average (at person-specific mean), and high ($+1SD$) levels of daily instrumental support.



(b) Trajectories of positive affect across low ($-1SD$), average, and high ($+1SD$) levels of mean instrumental support (over 22 days).

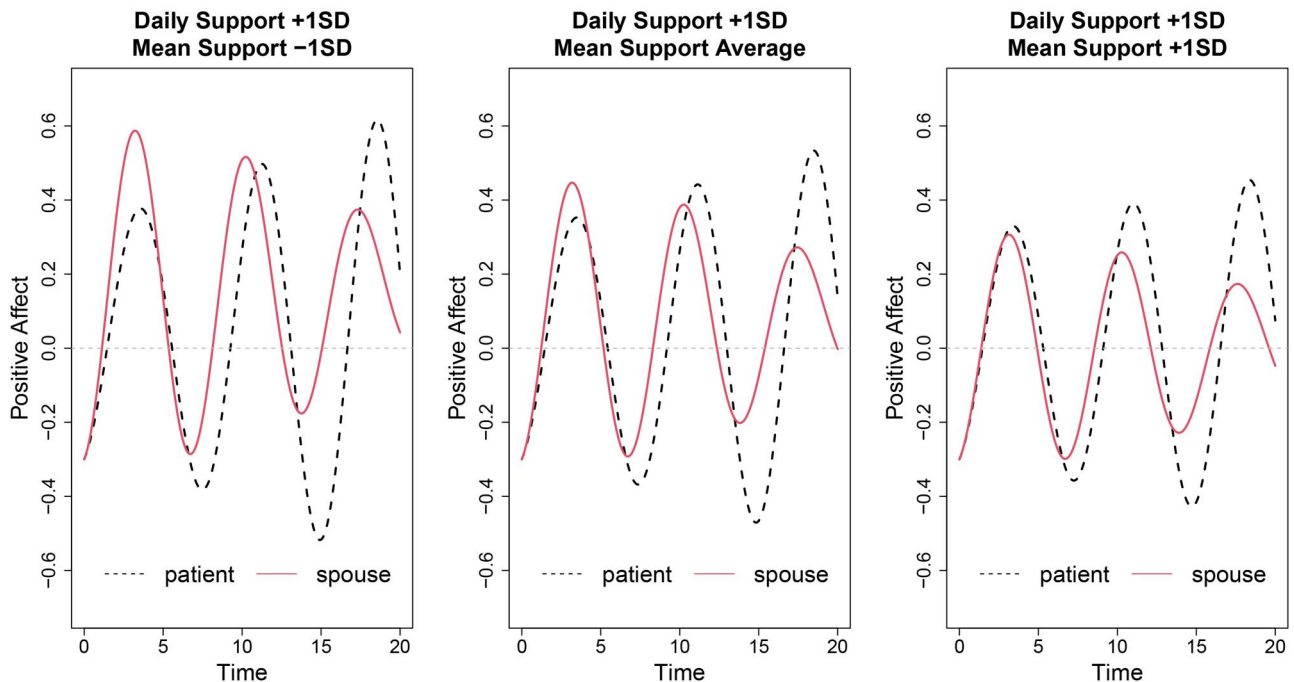


Figure 4. Simulated trajectories of a couple's positive affect.

Note. The black dashed line represents a patient's trend, and the red line represents a spouse's trend. Subfigure (a) shows that when spouses provide a higher-than-usual ($+1SD$) level of instrumental support over time, more extreme and positive upward shifts in their positive affect are expected over time. Under subfigure (b), the spouse's daily instrumental support provision was assumed to be consistently higher than usual ($+1SD$). Three trajectory plots in subfigure (b) indicate that the effect of providing higher-than-average daily support was much reduced for spouses whose typical level of instrumental support provision is high compared to those whose typical support level was low.

matrix as our input, resulting in six R^2 's for patients' positive affect: 0.323, 0.243, 0.214, 0.201, 0.232, and 0.324, with an average of 0.256. Therefore, approximately 25% of the level-1 variance in the spouse's positive affect was explained by the intercept and specified predictors, including both the fixed and random effects of the predictors. Regarding the effect of daily social support on a spouse's positive affect (β_{si}), the level-2 R^2 was 0.044, indicating that relatively little (approximately 4.4% of the) variance of the effect of daily social support was explained by its intercept and average social support.

In summary, we utilized a multilevel bivariate oscillator model structured as an M-LDSEM to investigate between-couple differences in coupling dynamics and the roles of instrumental support at the within- and between-person levels. Furthermore, applying the Bayesian framework to M-LDSEM allowed us to incorporate relatively high-dimensional random effects, while it did not converge when the same corresponding model was specified using the frequentist approach.

Discussion

We sought to address several key challenges of applying the M-LDSEM method to multivariate short time series, which are frequently observed in the behavioral sciences. We also expanded the usability of M-LDSEM by investigating the tenability of incorporating time- and person-specific covariates into differential equation models. Our simulation results demonstrated the benefits of the inclusion of higher-order derivatives and a Bayesian inferential method in parameter and SE estimation. In line with the previous study by Boker et al. (2020), we found that the 4th-order model is preferred in general to the 2nd-order model given that the improvement of accuracy for most parameters far outweighed the increased variability associated with the use of the 4th-order M-LDSEM compared to the conventional 2nd-order approach. However, we suggest employing the 2nd-order model for those focusing on obtaining an accurate estimate of the damping parameter.

Consonant with the expected improvements afforded by Bayesian methods in estimating complex models (i.e., analyzing nested data with differential equation model with random slopes), the Bayesian approach outperformed the frequentist estimations considered not only in terms of the convergence rate but also the properties of the point and SE estimates. The Bayesian approach adopted in the current study

utilized relatively uninformative priors that have been shown, in some special cases, to yield results that are similar to frequentist approaches. In addition, our inferential conclusions and conceptualization of the parameters in the data generation processes (e.g., specifying the true parameter values as fixed as opposed to random) also mirror common principles in the frequentist framework. Thus, key differences between the approaches compared were primarily in the estimation engines that underlie these approaches (e.g., sampling from the joint posterior distribution versus maximizing a likelihood). All of the approaches considered involved some form of approximation, but the specific Bayesian approach considered outperformed other frequentist alternatives in the proposed model. By adopting the Bayesian framework, we also have access to the empirical posterior distributions of modeling parameters, which offer more possibilities for quantifying the uncertainty of these components.

Moreover, in cases where there are clear reasons to suspect deviations from normality, the Bayesian framework provides the flexibility to accommodate non-Gaussian distributional assumptions (Berger et al., 1994). For instance, adopting alternative non-Gaussian prior distributions, such as heavy-tailed t distributions, may prove robust in handling violations of the independent residual assumption, especially in the context of using time delay embedded data. It is important to note that a thorough exploration of these robust variations could improve Bayesian estimation performance even further, although such in-depth analysis is beyond the scope of the present study and warrants additional investigation. We do acknowledge that frequentist approaches might provide much greater computational efficiency in cases involving simpler models, and appropriate prior choices also play a critical role in successful Bayesian estimation (Van Dongen, 2006).

In summary, through both empirical and simulation results, we have demonstrated that the Bayesian estimator with the 4th-order M-LDSEM can be applied to relatively short time series. However, there remain a few unresolved issues that warrant further evaluation in future studies. First, although power estimates remained consistently high, the best-performing combination still displayed notable biases in coupling parameter accuracy. This approach shows substantial improvements over conventional LDSEM results reported in the previous simulation study (Hu et al., 2014), despite the increased complexity of the models considered here, notably the incorporation of a multilevel extension and time-varying covariates.

Despite these developments not completely meeting the stringent criteria of empirical researchers, they contribute valuable insights in a field where simulation studies on the effects of coupling in differential equations are scarce. Additionally, our novel comparison between robust variants and the Bayesian method in a complex model with limited samples offers practical benefits, warranting future research to expand on our findings.

Our study highlighted the need to evaluate the consequences of data preparation procedures such as detrending when time-varying covariates are included. Researchers should carefully consider the functional form and theoretical aspect of the time-varying covariates when deciding whether detrending is necessary. If an existing time trend is not relevant to the core research interest and may lead to incorrect conclusions concerning the key processes of interest, detrending is recommended. However, if the trend over time is meaningful and could potentially explain the dynamics of interest (as in our case with the OU time-varying covariate), we do not recommend detrending but rather, would suggest direct inclusion of all relevant variables in the fitted model.

Another issue concerns ways of handling missing data. Differential equation models by nature posit the unfolding of the processes of interest continuously at any time point. Thus, missingness that arises in designs targeting equally spaced measurement intervals can just be incorporated seamlessly into differential equation models as unequally spaced time intervals (Oud & Voelkle, 2014). However, it still entails unnecessary data losses in M-LDSEM. It is because individual data with severe missingness are (or have to be) omitted from analyses if measured occasions are sparser than the embedding dimension. Thus, researchers should consider the cost and benefits of using the M-LDSEM approach when the total number of occasions is not long and missingness is severe. As other possibilities, we could insert missingness into the input data or utilize missing data handling methods in the time series (Li et al., 2019; Ji et al., 2020). Further investigation of missing data handling methods in the context of M-LDSEM models is an important methodological extension.

While our findings support the effectiveness of the 4th-order M-LDSEM plus Bayesian estimator with short time series, our simulation scope is limited. Assessing the viability of M-LDSEM methods with panel data (< 10 time points) is crucial, considering potential drawbacks in using time-delay embedding in such cases. Comparing with other continuous-time

methods, such as exact discrete (Oud & Jansen, 2000), successfully applied to panel data with as few as 4 time points (Oud, 2017), would be worthwhile. Furthermore, all the sample sizes considered in the present study were larger than the median sample size of $n=87$ reported in previous review studies Konjarski et al. (2018); Liu et al. (2019); Yang et al. (2019). A more thorough evaluation of the effects of using even smaller sample sizes is warranted. Additionally, the results of the simulation may differ under a different set of parameters, necessitating further studies to explore diverse parameter sets for valuable insights.

Another topic for future direction is investigating the optimal level of the highest derivative. Although using higher-order derivatives is likely to improve accuracy, it is important to note that excessive use of terms in the expansion may increase the computational burden and may not necessarily lead to a more precise estimation (e.g., damping parameter). Thus, this topic warrants further systematic investigation.

Further research on effect size measures for M-LDSEM is needed. While the R^2 measures available in Mplus can serve as computationally viable proxies for effect sizes, further work to derive effect size measures for M-LDSEM based on similar concepts of effect sizes for multi-level models Rights and Sterba (2020) could lead to more meaningfully decomposed effect size measures. Furthermore, as shown in the empirical results, the use of a time-delay embedded matrix as input data presents challenges that require additional investigation.

LDSEM methods were designed to handle ordinary differential equation models. They may not be appropriate for fitting stochastic differential equation models with process noises (McKee et al., 2020). To the extent that the assumptions of LDSEM methods are fulfilled, the implementation of LDSEM is relatively straightforward, computationally efficient, and allows users to leverage the advantages afforded by the SEM framework, such as ease of incorporating multiple measures and availability of fit indices. For instance, model fit could be assessed to determine the order of models.

Despite all the above limitations, our study provides a scaffold to explore the evolving relationships of variables of interest in a more natural way by proposing a feasible approach to fit a continuous-time model that works even in the context of the time series of the finite length and the multilevel extension. Moreover, our approach offers several practical advantages. Being seamlessly integrated into SEM, it readily

accommodates multivariate measures and incorporates the effects on antecedents and outcomes, enhancing the investigation of complex relationships. Additionally, it is computationally efficient compared to other continuous-time modeling methods (Driver et al., 2017; Ruissen et al., 2022), providing a faster alternative for model fitting, particularly useful for large datasets or intricate models. This will bridge the gap between methodology and reality, making continuous-time models with numerous advantages more accessible and practical in the field of behavioral science.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by the NIH Intensive Longitudinal Health Behavior Cooperative Agreement Program under U24AA027684, National Science Foundation grants IGE-1806874, and SES-1823633, the National Center for Advancing Translational Sciences under UL1TR002014-06, and the National Institutes of Health under R01 AG026010.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

References

- Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 411–434. https://doi.org/10.1207/s15328007sem1203_4
- Asparouhov, T., & Muthen, B. (2010). Bayesian analysis of latent variable models using mplus.
- Asparouhov, T., Muthen, B. (2006). Multilevel modeling of complex survey data. *Proceedings of the Joint Statistical Meeting in Seattle*, 2718–2726.
- Asparouhov, T., Muthen, B. (2007). Computationally efficient estimation of multilevel high-dimensional latent variable models. *Proceedings of the 2007 Joint Statistical Meeting in Salt Lake City, Utah, Section on Statistics in Epidemiology*, 2531–2535.
- Beach, S. R., Schulz, R., Yee, J. L., & Jackson, S. (2000). Negative and positive health effects of caring for a disabled spouse: Longitudinal findings from the caregiver health effects study. *Psychology and Aging*, 15(2), 259–271. <https://doi.org/10.1037/0882-7974.15.2.259>
- Berger, J. O., Moreno, E., Pericchi, L. R., Bayarri, M. J., Bernardo, J. M., Cano, J. A., De la Horra, J., Martín, J., Ríos-Insúa, D., Betrò, B., Dasgupta, A., Gustafson, P., Wasserman, L., Kadane, J. B., Srinivasan, C., Lavine, M., O'Hagan, A., Polasek, W., Robert, C. P., ... Sivaganesan, S. (1994). An overview of robust Bayesian analysis. *Test*, 3(1), 5–124. <https://doi.org/10.1007/BF02562676>
- Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 161–178). Routledge.
- Boker, S. M., & Graham, J. (1998). A dynamical systems analysis of adolescent substance abuse. *Multivariate Behavioral Research*, 33(4), 479–507. https://doi.org/10.1207/s15327906mbr3304_3
- Boker, S. M., Moulder, R. G., & Sjöbeck, G. R. (2020). Constrained fourth order latent differential equation reduces parameter estimation bias for damped linear oscillator models. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(2), 202–218. <https://doi.org/10.1080/10705511.2019.1641816>
- Boker, S. M., Neale, M., & Rausch, J. (2004). Latent differential equation modeling with multivariate multi-occasion indicators. In *Recent developments on structural equation models* (pp. 151–174). Springer.
- Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, 37(1), 127–160. https://doi.org/10.1207/S15327906MBR3701_06
- Bonanno, G. A. (2001). Emotion self-regulation.

- Butler, E. A. (2011). Temporal interpersonal emotion systems: The "ties" that form relationships. *Personality and Social Psychology Review*, 15(4), 367–393. <https://doi.org/10.1177/1088868311411164>
- Chen, M., Chow, S.-M., Oravecz, Z., & Ferrer, E. (2023). Fitting Bayesian stochastic differential equation models with mixed effects through a filtering approach. *Multivariate Behavioral Research*, 58(5), 1014–1038. <https://doi.org/10.1080/00273171.2023.2171354>
- Chow, S.-M. (2019). Practical tools and guidelines for exploring and fitting linear and nonlinear dynamical systems models. *Multivariate Behavioral Research*, 54(5), 690–718. <https://doi.org/10.1080/00273171.2019.1566050>
- Chow, S.-M., Haltigan, J. D., & Messinger, D. S. (2010). Dynamic infant–parent affect coupling during the face-to-face/still-face. *Emotion (Washington, D.C.)*, 10(1), 101–114. <https://doi.org/10.1037/a0017824>
- Chow, S.-M., Ho, M.-h R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modelling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(2), 303–332. <https://doi.org/10.1080/10705511003661553>
- Chow, S.-M., Lu, Z., Sherwood, A., & Zhu, H. (2016). Fitting nonlinear ordinary differential equation models with random effects and unknown initial conditions using the stochastic approximation expectation maximization (SAEM) algorithm. *Psychometrika*, 81(1), 102–134. <https://doi.org/10.1007/s11336-014-9431-z>
- Chow, S.-M., Ram, N., Boker, S. M., Fujita, F., & Clore, G. (2005). Emotion as a thermostat: Representing emotion regulation using a damped oscillator model. *Emotion (Washington, D.C.)*, 5(2), 208–225. <https://doi.org/10.1037/1528-3542.5.2.208>
- De Leeuw, J., De Graeff, A., Ros, W., Hordijk, G., Blijham, G., & Winnubst, J. (2000). Negative and positive influences of social support on depression in patients with head and neck cancer: A prospective study. *Psycho-Oncology*, 9(1), 20–28. [https://doi.org/10.1002/\(SICI\)1099-1611\(200001/02\)9:1<20::AID-PON425>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1099-1611(200001/02)9:1<20::AID-PON425>3.0.CO;2-Y)
- Depaoli, S., & Clifton, J. P. (2015). A Bayesian approach to multilevel structural equation modeling with continuous and dichotomous outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(3), 327–351. <https://doi.org/10.1080/10705511.2014.937849>
- Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with r package ctsem. *Journal of Statistical Software*, 77(5), 1–35. <https://doi.org/10.18637/jss.v077.i05>
- Durham, G. B., & Gallant, A. R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business & Economic Statistics*, 20(3), 297–338. <https://doi.org/10.1198/073500102288618397>
- Elerian, O., Chib, S., & Shephard, N. (2001). Likelihood inference for discretely observed nonlinear diffusions. *Econometrica*, 69(4), 959–993. <https://doi.org/10.1111/1468-0262.00226>
- Elfenbein, H. A. (2014). The many faces of emotional contagion: An affective process theory of affective linkage. *Organizational Psychology Review*, 4(4), 326–362. <https://doi.org/10.1177/2041386614542889>
- Feinberg, M. E., Xia, M., Fosco, G. M., Heyman, R. E., & Chow, S.-M. (2017). Dynamical systems modeling of couple interaction: A new method for assessing intervention impact across the transition to parenthood. *Prevention Science: The Official Journal of the Society for Prevention Research*, 18(8), 887–898. <https://doi.org/10.1007/s11121-017-0803-3>
- Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 88(3), 296–308. <https://doi.org/10.1016/j.ijpsycho.2012.10.013>
- Freedman, D. A. (2006). On the so-called "Huber sandwich estimator" and "robust standard errors". *The American Statistician*, 60(4), 299–302. <https://doi.org/10.1198/000313006X152207>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman; Hall/CRC.
- Gremore, T. M., Baucom, D. H., Porter, L. S., Kirby, J. S., Atkins, D. C., & Keefe, F. J. (2011). Stress buffering effects of daily spousal support on women's daily emotional and physical experiences in the context of breast cancer concerns. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 30(1), 20–30. <https://doi.org/10.1037/a0021798>
- Hallquist, M. N., & Wiley, J. F. (2018). Mplusautomation: An r package for facilitating large-scale latent variable analyses in m plus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 621–638. <https://doi.org/10.1080/10705511.2017.1402334>
- Hilpert, P., Brick, T. R., Flückiger, C., Vowels, M. J., Ceulemans, E., Kuppens, P., & Sels, L. (2020). What can be learned from couple research: Examining emotional co-regulation processes in face-to-face interactions. *Journal of Counseling Psychology*, 67(4), 475–487. <https://doi.org/10.1037/cou0000416>
- Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. *Oxford Handbook of Quantitative Methods*, 2(1), 281–294.
- Hu, Y., Boker, S. M., Neale, M., & Klump, K. L. (2014). Coupled latent differential equation with moderators: Simulation and application. *Psychological Methods*, 19(1), 56–71. <https://doi.org/10.1037/a0032476>
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 221–233). University of California Press.
- Ji, L., Chen, M., Oravecz, Z., Cummings, E. M., Lu, Z.-H., & Chow, S.-M. (2020). A Bayesian vector autoregressive model with nonignorable missingness in dependent variables and covariates: Development, evaluation, and application to family processes. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(3), 442–467. <https://doi.org/10.1080/10705511.2019.1623681>
- Katinka, H., Boker, S. M., & Bergeman, C. S. (2020). A note on the usefulness of constrained fourth-order latent differential equation models in the case of small t. *Psychometrika*, 85(4), 1016–1027. <https://doi.org/10.1007/s11336-020-09738-x>

- Konjarski, M., Murray, G., Lee, V. V., & Jackson, M. L. (2018). Reciprocal relationships between daily sleep and mood: A systematic review of naturalistic prospective studies. *Sleep Medicine Reviews*, 42, 47–58. <https://doi.org/10.1016/j.smrv.2018.05.005>
- Koval, P., Kuppens, P., Allen, N. B., & Sheeber, L. (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition & Emotion*, 26(8), 1412–1427. <https://doi.org/10.1080/02699931.2012.667392>
- Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological adjustment. *Psychological Science*, 21(7), 984–991. <https://doi.org/10.1177/0956797610372634>
- Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, 11(3), 129–141. https://doi.org/10.1207/S15327965PLI1103_01
- Li, Y., Ji, L., Oravecz, Z., Brick, T. R., Hunter, M. D., & Chow, S.-M. (2019). Dynr: mi: An r program for multiple imputation in dynamic modeling. *World Academy of Science, Engineering and Technology*, 13(5), 302.
- Liu, H., Xie, Q. W., & Lou, V. W. (2019). Everyday social interactions and intra-individual variability in affect: A systematic review and meta-analysis of ecological momentary assessment studies. *Motivation and Emotion*, 43(2), 339–353. <https://doi.org/10.1007/s11031-018-9735-x>
- Li, Y., Wood, J., Ji, L., Chow, S.-M., & Oravecz, Z. (2022). Fitting multilevel vector autoregressive models in stan, jags, and mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 29(3), 452–475. <https://doi.org/10.1080/10705511.2021.1911657>
- Lu, Z.-H., Chow, S.-M., Ram, N., & Cole, P. M. (2019). Zero-inflated regime-switching stochastic differential equation models for highly unbalanced multivariate, multi-subject time-series data. *Psychometrika*, 84(2), 611–645. <https://doi.org/10.1007/s11336-019-09664-7>
- Lu, Z.-H., Chow, S.-M., Sherwood, A., & Zhu, H. (2015). Bayesian analysis of ambulatory cardiovascular dynamics with application to irregularly spaced sparse data. *Annals of Applied Statistics*, 9(3), 1601–1620. <https://doi.org/10.1214/15-aos846>
- Marini, C. M., Wilson, S. J., Tate, A. M., Martire, L. M., & Franks, M. M. (2021). Short-and long-term effects of support visibility on support providers' negative affect. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 76(3), 461–470. <https://doi.org/10.1093/geronb/gbz114>
- Martire, L. M., Stephens, M. A. P., Mogle, J., Schulz, R., Brach, J., & Keefe, F. J. (2013). Daily spousal influence on physical activity in knee osteoarthritis. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 45(2), 213–223. <https://doi.org/10.1007/s12160-012-9442-x>
- Mbalawata, I. S., Särkkä, S., & Haario, H. (2013). Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering. *Computational Statistics*, 28(3), 1195–1223. <https://doi.org/10.1007/s00180-012-0352-y>
- McGlynn, F. D., Mealiea, W. L., Jr., & Landau, D. L. (1981). The current status of systematic desensitization. *Clinical Psychology Review*, 1(2), 149–179. [https://doi.org/10.1016/0272-7358\(81\)90001-5](https://doi.org/10.1016/0272-7358(81)90001-5)
- McKee, K. L., Hunter, M. D., & Neale, M. C. (2020). A method of correcting estimation failure in latent differential equations with comparisons to kalman filtering. *Multivariate Behavioral Research*, 55(3), 405–424. <https://doi.org/10.1080/00273171.2019.1642730>
- Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, 50(2), 181–202. <https://doi.org/10.1007/BF02294246>
- Monin, J. K., & Schulz, R. (2009). Interpersonal effects of suffering in older adult caregiving relationships. *Psychology and Aging*, 24(3), 681–695. <https://doi.org/10.1037/a0016355>
- Muthen, L., & Muthen, B. (2017). Mplus user's guide. Eighth version. Los Angeles, CA: Muthen & Muthen, 10, 1600–0447.
- Newton-John, T. R. (2013). How significant is the significant other in patient coping in chronic pain? *Pain Management*, 3(6), 485–493. <https://doi.org/10.2217/pmt.13.52>
- Niven, K. (2017). The four key characteristics of interpersonal emotion regulation. *Current Opinion in Psychology*, 17, 89–93. <https://doi.org/10.1016/j.copsyc.2017.06.015>
- 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>
- Oud, J. H. (2017). Comparison of four procedures to estimate the damped linear differential oscillator for panel data. In *Longitudinal models in the behavioral and related sciences* (pp. 19–39) Routledge.
- Oud, J. H., & Jansen, R. A. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, 65(2), 199–215. <https://doi.org/10.1007/BF02294374>
- Oud, J. H., & Voelkle, M. C. (2014). Do missing values exist? Incomplete data handling in cross-national longitudinal studies by means of continuous time modeling. *Quality & Quantity*, 48(6), 3271–3288. <https://doi.org/10.1007/s11135-013-9955-9>
- Poulin, M. J., Brown, S. L., Ubel, P. A., Smith, D. M., Jankovic, A., & Langa, K. M. (2010). Does a helping hand mean a heavy heart? Helping behavior and well-being among spouse caregivers. *Psychology and Aging*, 25(1), 108–117. <https://doi.org/10.1037/a0018064>
- R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reinhardt, J. P., Boerner, K., & Horowitz, A. (2006). Good to have but not to use: Differential impact of perceived and received support on well-being. *Journal of Social and Personal Relationships*, 23(1), 117–129. <https://doi.org/10.1177/0265407506060182>
- Rights, J. D., & Sterba, S. K. (2020). New recommendations on the use of r-squared differences in multilevel model comparisons. *Multivariate Behavioral Research*, 55(4), 568–599. <https://doi.org/10.1080/00273171.2019.1660605>
- Rodríguez-Blanco, L., Carballo, J. J., & Baca-García, E. (2018). Use of ecological momentary assessment (EMA) in non-suicidal self-injury (NSSI): A systematic review. *Psychiatry Research*, 263, 212–219. <https://doi.org/10.1016/j.psychres.2018.02.051>

- Ruissen, G. R., Beauchamp, M. R., Puterman, E., Zumbo, B. D., Rhodes, R. E., Hives, B. A., Sharpe, B. M., Vega, J., Low, C. A., & Wright, A. G. (2022). Continuous-time modeling of the bidirectional relationship between incidental affect and physical activity. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 56(12), 1284–1299. <https://doi.org/10.1093/abm/kaac024>
- Sanders, S., & Power, J. (2009). Roles, responsibilities, and relationships among older husbands caring for wives with progressive dementia and other chronic conditions. *Health & Social Work*, 34(1), 41–51. <https://doi.org/10.1093/hsw/34.1.41>
- Sbarra, D. A., & Hazan, C. (2008). Coregulation, dysregulation, self-regulation: An integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc.*, 12(2), 141–167. <https://doi.org/10.1177/1088868308315702>
- Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 495–515. <https://doi.org/10.1080/10705511.2017.1392862>
- 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>
- Sels, L., Ceulemans, E., & Kuppens, P. (2018). A general framework for capturing interpersonal emotion dynamics: Associations with psychological and relational adjustment. In A. K. Randall & D. Schoebi (Eds.), *Interpersonal emotion dynamics in close relationships* (1st ed., pp. 27–46) Cambridge University Press.
- Soetaert, K., Petzoldt, T., & Setzer, R. W. (2010). Solving differential equations in r: Package desolve. *Journal of Statistical Software*, 33(9), 1–25. <https://doi.org/10.18637/jss.v033.i09>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(3), 485–493. <https://doi.org/10.1111/rssb.12062>
- Steele, J. S., & Ferrer, E. (2011). Latent differential equation modeling of self-regulatory and coregulatory affective processes. *Multivariate Behavioral Research*, 46(6), 956–984. <https://doi.org/10.1080/00273171.2011.625305>
- Van Dongen, S. (2006). Prior specification in Bayesian statistics: Three cautionary tales. *Journal of Theoretical Biology*, 242(1), 90–100. <https://doi.org/10.1016/j.jtbi.2006.02.002>
- von Oertzen, T., & Boker, S. M. (2010). Time delay embedding increases estimation precision of models of intraindividual variability [Num Pages: 158-175 Place: Williamsburg, Netherlands Publisher: Springer Nature B.V.]. *Psychometrika*, 75(1), 158–175. <https://doi.org/10.1007/S11336-009-9137-9>
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838. <https://doi.org/10.2307/1912934>
- White, H. (1994). *Estimation, inference and specification analysis*. Cambridge University Press.
- Yang, Y. S., Ryu, G. W., & Choi, M. (2019). Methodological strategies for ecological momentary assessment to evaluate mood and stress in adult patients using mobile phones: Systematic review. *JMIR mHealth and uHealth*, 7(4), e11215. <https://doi.org/10.2196/11215>
- Zhang, Z., Hamagami, F., Lijuan Wang, L., Nesselroade, J. R., & Grimm, K. J. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, 31(4), 374–383. <https://doi.org/10.1177/0165025407077764>
- Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling: A Multidisciplinary Journal*, 15(3), 377–402. <https://doi.org/10.1080/10705510802154281>
- Zhang, Z., & Nesselroade, J. R. (2007). Bayesian estimation of categorical dynamic factor models. *Multivariate Behavioral Research*, 42(4), 729–756. <https://doi.org/10.1080/00273170701715998>
- Zill, D. G. (1993). *A first course in differential equations* (5th ed.). PWS-KENT Publishing Company.

Appendix A. Additional simulation with time-varying covariates following sine functions

In the current simulation study, the simulated time-varying covariate, which follows the Ornstein-Uhlenbeck (OU) process, exhibits a linear time trend. In our preliminary analysis, we found that detrending, in this case, could lead to the inadvertent removal of some effects of time-varying covariates, leading to poor estimation of parameters related to these covariates.

To address concerns about the potential effects of detrending on estimation results, we conducted an additional Monte Carlo simulation study in which the time-varying covariate followed a sine function, rather than a linear function of time. Specifically, we generated 500 replications with $T = 100$ and $n = 150$. We used the proposed Bayesian 4th-order M-LDSEM approach to fit our simulation model to the data. In this case, the removal of a linear time trend did not lead to biases in estimations, as shown in Table A1.

Table A1. Results with time-varying covariates following sine functions.

θ	Mean $\hat{\theta}$	rBias	RMSE	MCSD	$\widehat{aSE}$	RDSE	
Result without detrending							
η	-0.5	-0.514	0.028	0.022	0.0169	0.0096	-0.429
ζ	-0.04	-0.027	-0.331	0.014	0.0044	0.0062	0.3986
γ	-0.1	-0.078	-0.221	0.023	0.0058	0.0042	-0.2804
β_0	0.3	0.313	0.045	0.025	0.0207	0.0195	-0.0584
β_1	0.1	0.104	0.042	0.012	0.0117	0.0112	-0.0453
Result with detrending							
η	-0.5	-0.505	0.011	0.016	0.015	0.0093	-0.3812
ζ	-0.04	-0.027	-0.333	0.014	0.0042	0.0064	0.5141
γ	-0.1	-0.075	-0.249	0.026	0.006	0.0042	-0.3017
β_0	0.3	0.309	0.029	0.023	0.0209	0.0192	-0.0783
β_1	0.1	0.101	0.014	0.012	0.0115	0.011	-0.0366

Appendix B. Additional simulation with empirical model

There were gaps between our main simulation and the model used in our empirical analysis. To closely resemble the complexity of the empirical model presented in Equation 4, this simulation incorporated individual variability in the coupling parameter and used the damping parameter fixed at zero. We generated 500 replications with $T = 100$ and $n = 150$. We used the proposed Bayesian 4th-order M-LDSEM approach to fit our simulation model to the data. The result is shown in Table B1.

Consistent with our main simulations, we observed an underestimation of the coupling parameter and its related covariate effect. As anticipated, introducing individual differences in the coupling parameters made the estimation slightly more challenging. Thus, more caution is needed to interpret the coupling parameters when they are allowed to differ across participants. However, the directionality of the effects remained accurate, and the power was consistently at 100% for all parameters. Despite the integration of random effects and the increased model complexity, the recovery of the parameters related to time-varying covariates remained commendable. Hence, while our method may not be flawless, it still provides valuable insights, particularly when considering its computational efficiency.

Table B1. Results of the fixed parameters from the empirical model.

	θ	Mean $\hat{\theta}$	rBias	RMSE	MCSD	$a\widehat{SE}$	RDSE
η	-0.5	-0.457	-0.085	0.043	0.0036	0.0014	-0.6172
γ_0	-0.1	-0.067	-0.328	0.034	0.0067	0.0042	0.1595
γ_1	0.1	0.072	-0.285	0.029	0.0045	0.0042	0.1938
β_0	0.3	0.285	-0.050	0.023	0.0180	0.0170	-0.0513
β_1	0.1	0.098	-0.023	0.010	0.0099	0.0097	-0.0219