





THEORY AND METHODS

Enhancing Two-Stage Estimation in Differential Equation Models: A Bias-Correction Method via Stochastic Approximation

Xiaohui Luo¹ , Hongyun Liu², Yueqin Hu¹ and Yang Liu³ 

¹Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, China; ²Research Center for Capacity Building in Educational Assessment and Evaluation (Beijing Higher Education Innovation Center for Philosophy and Social Sciences), China; ³Department of Human Development and Quantitative Methodology, University of Maryland at College Park, USA

Corresponding authors: Yang Liu and Yueqin Hu; Emails: yliu87@umd.edu, yueqinhu@bnu.edu.cn

(Received 28 May 2025; revised 5 March 2026; accepted 6 March 2026)

Abstract

Differential equation models have become increasingly popular for investigating dynamic processes. However, commonly used two-stage estimation methods, such as the generalized local linear approximation (GLLA), often produce biased parameter estimates. This study proposes a bias-correction method for GLLA estimates in second-order differential equation models. The method solves a bias-correction equation (derived from the relation between true parameter values and their biased estimates) via stochastic approximation, producing asymptotically unbiased estimates even with large initial bias. We first demonstrate the application of the bias-correction method by correcting the bias of a single parameter in a simple second-order differential equation model (i.e., the linear oscillator model with time-independent measurement error). We then extend the method to a more commonly used second-order differential equation model (i.e., the damped linear oscillator model), examining its performance in simultaneously addressing multiple parameters and incorporating time-dependent dynamic error. A simulation study shows that the bias-correction method substantially reduces bias in GLLA estimates, yielding highly accurate and precise parameter estimates. An empirical illustration further compares the results of GLLA and the bias-correction method. Our findings highlight the effectiveness of the proposed method in improving parameter estimation for differential equation models, offering an enhanced approach for analyzing dynamic processes.

Keywords: bias correction; differential equation model; stochastic approximation; two-stage estimation

Differential equation models have become increasingly popular for investigating dynamic processes in social and behavioral sciences. These models describe the relations between a variable's value and its derivatives over time (Boker et al., 2004), offering a powerful framework for modeling the mechanisms underlying behavioral processes using intensive longitudinal data. In psychological research, second-order differential equation models, such as the damped oscillator model, are commonly employed to capture the dynamics of systems around its equilibrium. These models are especially useful for describing self-regulatory processes, which are central to many psychological theories (Boker et al., 2004; Hu & Treinen, 2019). Given their ability to model complex dynamics and their theoretical relevance, second-order differential equation models have been applied to examine dynamic mechanisms across a variety

© The Author(s), 2026. Published by Cambridge University Press on behalf of Psychometric Society.

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

ID	X_1	X_2	X_3	X_4	X_5
1	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$
1	$X_{1,2}$	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$	$X_{1,6}$
1	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$	$X_{1,6}$	$X_{1,7}$

2	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$X_{2,4}$	$X_{2,5}$
2	$X_{2,2}$	$X_{2,3}$	$X_{2,4}$	$X_{2,5}$	$X_{2,6}$
2	$X_{2,3}$	$X_{2,4}$	$X_{2,5}$	$X_{2,6}$	$X_{2,7}$

3	$X_{3,1}$	$X_{3,2}$	$X_{3,3}$	$X_{3,4}$	$X_{3,5}$
3	$X_{3,2}$	$X_{3,3}$	$X_{3,4}$	$X_{2,5}$	$X_{3,6}$
3	$X_{3,3}$	$X_{3,4}$	$X_{3,5}$	$X_{3,6}$	$X_{3,7}$

Figure 1. Illustration of a five-dimensional time-delay embedded data matrix ($N = 3$ subjects, $T = 7$ observations).

of psychological constructs, such as affects (Chow et al., 2005; Ollero et al., 2025; Reed et al., 2015; Steele & Ferrer, 2011), physiological reactivity (Helm et al., 2012), substance use (Boker & Graham, 1998), working memory (Gasimova et al., 2014), sexual desire (Farr et al., 2014), and parent–adolescent interactions (Zheng et al., 2024).

With the growing application of differential equation models, their estimation methods have seen substantial developments in recent decades. These methods generally fall into two categories: one-stage and two-stage methods. The one-stage method employs the analytic solution or numerical integration of the differential equation model to perform nonlinear least squares regression. For instance, Hu and Treinen (2019) demonstrated high estimation accuracy for parameters and their standard errors (SEs) in a simulation study based on the analytic solution of the damped oscillator model. However, this method requires multiple starting values (for the variable, its derivatives and model parameters), and the estimation results can be highly sensitive to these starting values, as will be illustrated in Section 1. Another one-stage approach is based on state-space models (Durbin & Koopman, 2012), which provide a general modeling framework for dynamic processes, including those governed by differential equations. These models are often regarded as an effective approach for estimating differential equation parameters, especially given their ability to handle measurement and dynamic errors within a unified framework. This approach can be implemented in R packages, such as *dynr* (Ou et al., 2019), *OpenMx* (Neale et al., 2016), and *ctsem* (Driver et al., 2017).

The two-stage methods are grounded in the linear relations between a variable and its derivatives in differential equation models. These methods first estimate derivatives from the observed data, and then use these estimated derivatives to obtain model parameters through linear regression. This procedure offers a practical advantage, as it can be easily implemented using standard linear regression tools. A widely used method in this category is the generalized local linear approximation (GLLA; Boker et al., 2010), which computes derivatives from observed data and estimates parameters in a linear regression model. Before estimating derivatives and model parameters, an essential preprocessing step is to reorganize the dataset into a time-delay embedded matrix. Specifically, the time series is cut into overlapping segments, which are then stacked as rows in the matrix. The length of these overlapping segments, referred to as the embedding dimension, determines the number of local time points used

for derivative estimation. An example of a five-dimensional time-delay embedded matrix is shown in Figure 1 ($N = 3$ subjects, $T = 7$ observations). This time-delay embedding technique enhances the precision of parameter estimation using overlapping samples (Von Oertzen & Boker, 2010) and has been shown to be robust to misspecification of sampling intervals (Boker et al., 2019).

The latent differential equation (LDE; Boker et al., 2004) method is a variant of GLLA that also builds upon the time-delay embedded data matrix. Unlike GLLA, which explicitly calculates the derivatives and treats them as observed variables in differential equation models, LDE specifies derivatives as latent variables indicated by observed time-delay embedded variables with fixed factor loadings.¹ Coefficients of the differential equation model are estimated as path coefficients among latent derivative variables. Although LDE can leverage the advantages of the SEM framework, both GLLA and LDE may produce biased estimates of differential equation coefficients at the second stage due to potential inaccuracies in derivative estimation at the first stage—whether through explicit calculation in GLLA or fixed loading specification in LDE. These limitations of GLLA and LDE will be further demonstrated in Section 1.

Given the flexibility, ease of implementation, and wide applications of GLLA in estimating differential equation models, addressing the bias in its parameters estimates is of critical importance. To this end, we propose a bias-correction method for the GLLA estimates in second-order differential equation models. The main idea behind this method is solving a bias-correction equation (derived from the relation between the true parameter values and their biased estimates) via stochastic approximation (SA). We demonstrate that this approach yields asymptotically unbiased point estimators even when the initial estimator (i.e., GLLA in our case) is asymptotically biased.² Additionally, we develop a numerical procedure to compute asymptotically valid SEs for the bias-corrected estimator.

The remainder of the article is organized as follows. We begin by illustrating the limitations of existing estimation methods, including the one-stage nonlinear regression approach and the two-stage methods (GLLA and LDE), using a simulated dataset based on a simple second-order differential equation model (i.e., the linear oscillator model with time-independent measurement error). Next, we introduce the bias-correction method for two-stage GLLA estimates, outlining its general idea and implementation via SA, and demonstrate its application to correct the bias of a single parameter using the same simulated dataset. We also provide a numerical computation procedure of the SEs for the bias-corrected estimator. We then extend the bias-correction method to a more commonly used second-order differential equation model—the damped linear oscillator model. In this extension, we first examine the performance of the bias-correction method in simultaneously addressing multiple parameters and then further investigate its effectiveness under more realistic scenarios by incorporating time-dependent dynamic error. Following this, we conduct a simulation study to evaluate the performance of the bias-correction method in handling multiple parameters in the damped linear oscillator model with dynamic error. Throughout these extended analyses and the simulation study, we compare the performance of our bias-correction method against estimates obtained from state-space models. We further use empirical data to demonstrate different results of GLLA and the bias-correction method. Finally, we discuss the broader implications of this study and suggest directions for future research.

1. Limitations of existing estimation methods: A simple illustration

The limitations of the one-stage method (i.e., nonlinear regression estimation) and the two-stage methods (GLLA and LDE) are illustrated using the linear oscillator model with measurement error.

¹Some researchers may regard LDE as a one-stage method because the estimation is conducted in a single step within the SEM framework. In this article, we categorize it as a two-stage method to highlight that its input data are not the raw observed time series but a time-delay embedded matrix pre-processed from the observations. This distinction allows us to contrast LDE with more typical one-stage methods that estimate parameters directly from the raw time-series data.

²In this article, we use the term estimator in its general sense, referring to a rule for calculating parameter estimates from observed data. The proposed bias-corrected estimator is asymptotically unbiased when both the number of individuals (N) and time points (T) tend to infinity.

The model is specified as

$$X(t) = x(t) + e(t), \quad (1)$$

$$\ddot{x}(t) = \eta x(t), \quad (2)$$

where $X(t)$ is the measured displacement of a variable from its equilibrium at time t , $x(t)$ is the latent value of the displacement at time t , $e(t)$ is the measurement error that follows a normal distribution $e(t) \sim N(0, \sigma_e)$, $\ddot{x}(t)$ is the second derivative of $x(t)$ with respect to time at time t , and η is the frequency parameter that determines the frequency of oscillation. η is restricted to be negative. Equations (1) and (2) describe a self-regulatory system: when the system deviates from its equilibrium, it is pulled back to its equilibrium. Higher absolute values of η indicate more frequent fluctuations.³

1.1. Data generation

A dataset with a large number of time points ($T = 1,000$) for a single individual ($N = 1$) was generated according to Equations (1) and (2). Specifically, we first simulated data without measurement error using the analytic solution of Equation (2):

$$x(t) = A \cos(\omega t) + B \sin(\omega t), \quad (3)$$

where $\omega = \sqrt{-\eta}$, $A = x_1$, and $B = \dot{x}_1/\omega$. Here, x_1 and \dot{x}_1 are the initial values of the displacement and the first derivative, respectively. We set $\eta = -0.8$ to account for a typical cycle length of 7 time points (i.e., $2\pi/\omega = 2\pi/\sqrt{-\eta} \approx 7$), which, for instance, corresponds to a weekly cycle with one observation per day.

1.2. One-stage estimation

In the one-stage method (i.e., nonlinear regression estimation), we used the analytic solution (Equation (3)) of the linear oscillator model to perform a nonlinear least square regression, aiming to estimate η , x_1 , and \dot{x}_1 by minimizing the residual sum of squares (RSS). This was implemented using the R function `nls`.

However, we found that the estimates were highly sensitive to the choice of starting values, particularly for η . As shown in Figure 2, the RSS varies as a function of η when x_1 and \dot{x}_1 are fixed at their true values. The RSS function has multiple local minima; therefore, numerical search algorithms may get trapped at different suboptimal solutions depending on the user-supplied starting values of η . The presence of local optima poses computational challenges for one-stage estimation of differential equation models—even when the model can be expressed in closed form.

1.3. Two-stage estimation

In the two-stage methods (GLLA and LDE), we reorganized the dataset to construct a time-delay embedded matrix with an embedding dimension of 5 (as illustrated in Figure 1). For GLLA, we first estimated the zero- and second-order derivatives using this matrix, and then fitted a linear regression model using the R function `lm` to estimate the regression effect (i.e., η) of the zero-order derivative $x(t)$ on the second-order derivative $\ddot{x}(t)$. For LDE, we specified a structural equation model using the R package `lavaan` (Rosseel, 2012), treating the derivatives as latent variables, the dimensions of the time-delay embedded data as observed variables, and fixing the factor loadings to predefined values (Boker *et al.*, 2004; Hu *et al.*, 2014).

The results showed that both methods produced biased estimates of η , with GLLA yielding an estimate of -0.618 and LDE yielding -0.626 , both deviating substantially from the true value of -0.8 .

³Example latent trajectories generated from this model, as well as from the damped oscillator models introduced later, are provided in Section S1 of the Supplementary Material for visual reference.

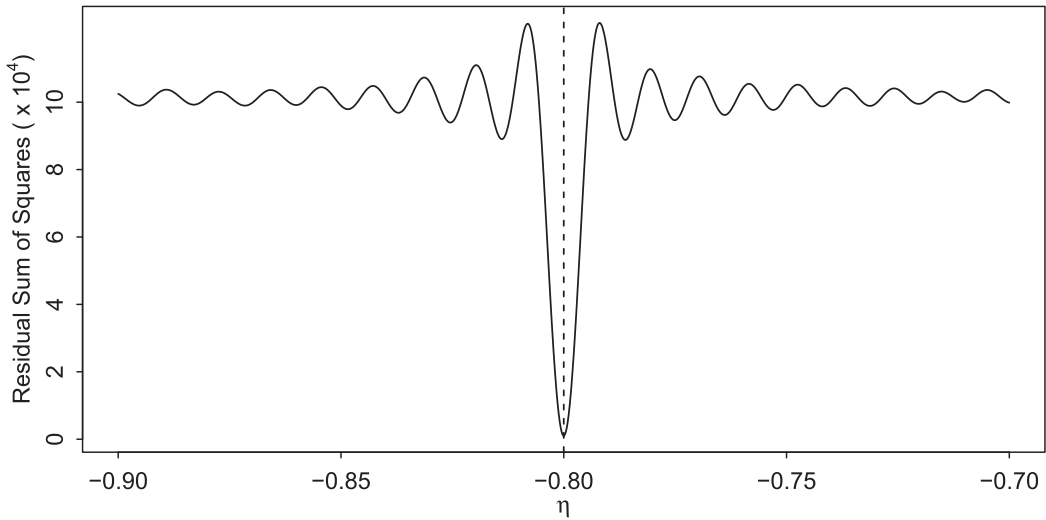


Figure 2. Residual sum of squares (RSS) across different values of η in the linear oscillator model.
 Note: The solid line represents the RSS for each candidate η . The vertical dashed line indicates the true value of η used in data generation.

These findings highlight the limitations of the two-stage methods in providing accurate estimates of the dynamic parameter η in differential equation models.⁴

2. Bias correction for two-stage GLLA estimates

In this section, we demonstrate how to correct the bias of two-stage GLLA estimates in differential equation models. We first introduce the general idea of the bias-correction method, and its implementation via SA. Then, the same dataset generated based on the linear oscillator model was used to illustrate the procedure of the bias-correction method.

2.1. Bias correction via stochastic approximation

2.1.1. General idea of the bias-correction method

Let $\mathbf{Y} = (Y_1, \dots, Y_n)'$ represent sample data generated from the model \mathbb{P}_{θ^*} , where $\theta^* \in \Theta$ is the true parameter vector, and $\hat{\theta} = \hat{\theta}(\mathbf{Y})$ is a point estimate of θ^* .⁵ The bias of $\hat{\theta}$ can be expressed as

$$b(\theta^*) = \mathbb{E}_{\theta^*} \{ \hat{\theta}(\mathbf{Y}) \} - \theta^*. \tag{4}$$

If there exists a θ such that $\mathbb{E}_{\theta} \{ \hat{\theta}(\mathbf{Y}) \} = \mathbb{E}_{\theta^*} \{ \hat{\theta}(\mathbf{Y}) \}$, and the map from θ to $\mathbb{E}_{\theta} \{ \hat{\theta}(\mathbf{Y}) \}$ is local injective (i.e., $\mathbb{E}_{\theta_1} \{ \hat{\theta}(\mathbf{Y}) \} = \mathbb{E}_{\theta_2} \{ \hat{\theta}(\mathbf{Y}) \} \Rightarrow \theta_1 = \theta_2$), we know that this θ must be the true value θ^* . However, since $\mathbb{E}_{\theta^*} \{ \hat{\theta}(\mathbf{Y}) \}$ is unknown, we approximate it with its unbiased estimate $\hat{\theta}(\mathbf{y})$, where $\mathbf{y} = (y_1, \dots, y_n)'$ denotes a specific realization (or observed counterpart) of the random sample \mathbf{Y} . Then, if we can find a θ such that $\mathbb{E}_{\theta} \{ \hat{\theta}(\mathbf{Y}) \} = \hat{\theta}(\mathbf{y})$, we can infer that this θ should be close to θ^* . Thus, the goal of bias

⁴The magnitude of η influences the accuracy of two-stage estimates. For instance, when $\eta = -0.2$ (corresponding to a cycle length of about 14 time points), the bias of GLLA estimates is notably smaller (GLLA estimate = -0.190). In this case, the bias-correction method introduced below further improves estimation accuracy (bias-corrected estimate = -0.200). Please see Section S2 of the Supplementary Material for details.

⁵Vectors and matrices are distinguished from scalars by boldface in our notation. This notational convention is followed throughout the article.

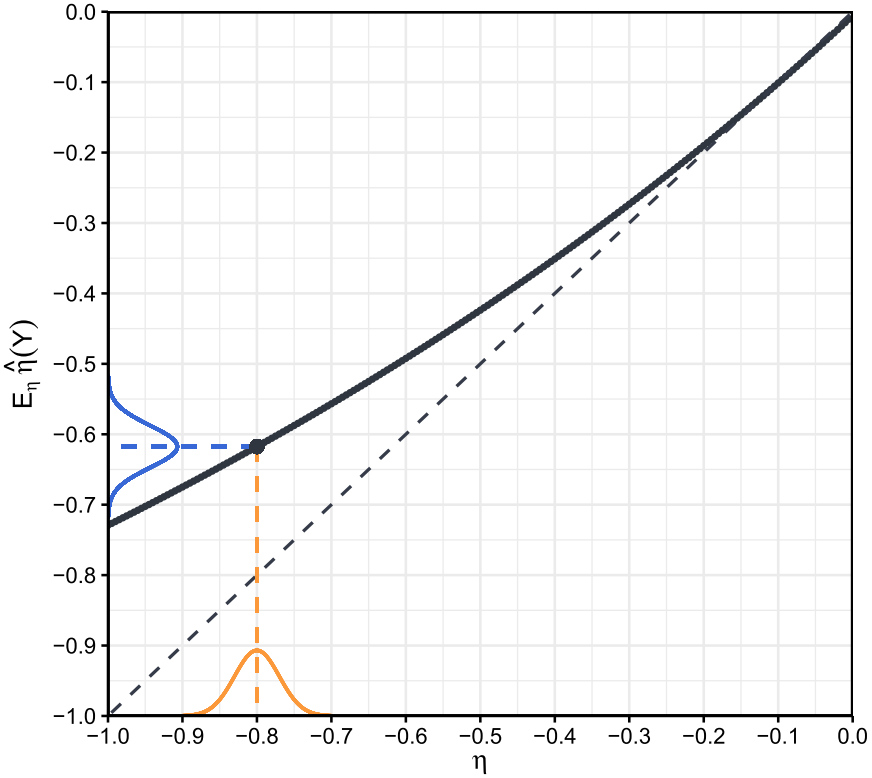


Figure 3. Illustration of the bias-correction procedure using the η estimation in the linear oscillator model as an example. *Note:* The solid black line represents the expectation $\mathbb{E}_{\eta}\{\hat{\eta}(\mathbf{Y})\}$ as a function of η . The orange dashed line marks the true value of $\eta = -0.8$, and the blue and orange distributions show the sampling distributions of the biased estimator $\hat{\eta}(\mathbf{Y})$ and the bias-corrected estimator $\tilde{\eta}(\mathbf{y})$, respectively.

correction becomes solving for θ from the following equation:

$$\mathbb{E}_{\theta}\{\hat{\theta}(\mathbf{Y})\} - \hat{\theta}(\mathbf{y}) = 0. \tag{5}$$

The solution to (5), denoted by $\tilde{\theta}(\mathbf{y})$, has been shown by Leung and Wang (1998) to reduce the bias to $O(n^{-1})$ for estimating single-parameter models under mild assumptions. In the Appendix, we generalize the bias-correction result in multiparameter settings, demonstrating that the bias-corrected estimator maintains the same $O(n^{-1})$ bias order and thus remains asymptotically unbiased.

The rationale behind this approach is illustrated in Figure 3, using the estimation of η in the linear oscillator model as an example. We varied the true values of η from -1 to 0 (excluding 0), and fixed the other parameters at the values used in the previous illustrative example to generated datasets. For each value of η , 10,000 datasets were generated, and η was estimated using GLLA. The expectation of η (denoted by $\mathbb{E}_{\eta}\{\hat{\eta}(\mathbf{Y})\}$) is plotted as a function of η in Figure 3 (the solid black line). The estimator $\hat{\eta}(\mathbf{Y})$ is found to be biased, and the bias increases as the magnitude of η increases. This trend is reflected in the deviation of $\mathbb{E}_{\eta}\{\hat{\eta}(\mathbf{Y})\}$ from the diagonal (the dotted black line). For the true value of $\eta = -0.8$ (indicated by the orange dashed line), GLLA overestimated the parameter (i.e., $\mathbb{E}_{-0.8}\{\hat{\eta}(\mathbf{Y})\} > -0.8$). To correct the bias, we can use the biased estimator $\hat{\eta}(\mathbf{Y})$, whose sampling distribution is centered around $\mathbb{E}_{-0.8}\{\hat{\eta}(\mathbf{Y})\}$ (depicted by the blue distribution), to obtain the sampling distribution for $\tilde{\eta}(\mathbf{y})$ (the orange distribution) by identifying the corresponding x -coordinate from the solid black line. The sampling distribution of the bias-corrected estimator $\tilde{\eta}(\mathbf{y})$ is expected to center around the true value of $\eta = -0.8$.

2.1.2. Stochastic approximation and the Robbins–Monro algorithm

To solve Equation (5), SA methods can be used to obtain bias-corrected estimates. These methods are particularly well-suited for solving equations involving intractable expectations. Specifically, let \mathbf{A} be a data-generating algorithm that produces simulated data \mathbf{Y} based on the parameter $\boldsymbol{\theta}$ and a random variable \mathbf{U} with distribution \mathbb{P} independent of $\boldsymbol{\theta}$, that is, $\mathbf{Y} = \mathbf{A}(\boldsymbol{\theta}, \mathbf{U})$. Consider solving for $\boldsymbol{\theta}$ from the equation (or a set of equations)

$$0 = \mathbf{g}(\boldsymbol{\theta}) = \mathbb{E}\{\mathbf{G}(\boldsymbol{\theta}, \mathbf{U})\}, \tag{6}$$

where $\mathbf{g}(\boldsymbol{\theta}) = \mathbb{E}\{\hat{\boldsymbol{\theta}}[\mathbf{A}(\boldsymbol{\theta}, \mathbf{U})]\} - \hat{\boldsymbol{\theta}}(\mathbf{y})$ and $\mathbf{G}(\boldsymbol{\theta}, \mathbf{U}) = \hat{\boldsymbol{\theta}}[\mathbf{A}(\boldsymbol{\theta}, \mathbf{U})] - \hat{\boldsymbol{\theta}}(\mathbf{y})$. The expectation is taken over the random variable $\mathbf{U} \sim \mathbb{P}$ while the observed data \mathbf{y} are kept fixed. In the previously discussed example based on the linear oscillator model, the frequency parameter η can be viewed as the parameter $\boldsymbol{\theta}$, and the measurement error $e(t)$ can be interpreted as parameter-free random components \mathbf{U} .⁶

To solve Equation (6), a fundamental SA method—the Robbins–Monro algorithm (RM algorithm; Robbins & Monro, 1951)—can be applied. This algorithm iteratively updates the approximation to the root using the following recursive scheme:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \gamma_k \hat{\mathbf{g}}_m(\boldsymbol{\theta}_k), \tag{7}$$

where $\boldsymbol{\theta}_k$ denotes the k -th iterate ($k = 1, 2, \dots, K$; K is the total number of iterations) of $\boldsymbol{\theta}$, γ_k is a sequence of constants converging to zero, and $\hat{\mathbf{g}}_m(\boldsymbol{\theta}_k)$ is an unbiased Monte Carlo (MC) estimate of $\mathbf{g}(\boldsymbol{\theta}_k)$, which is computed as

$$\hat{\mathbf{g}}_m(\boldsymbol{\theta}_k) = \frac{1}{m} \sum_{j=1}^m \mathbf{G}(\boldsymbol{\theta}_k, \mathbf{U}_j). \tag{8}$$

Here, $\mathbf{U}_1, \dots, \mathbf{U}_m$ are independent and identically distributed samples drawn from \mathbb{P} , and m is the total number of MC simulations (which is typically not large in practice). In this study, we set $m = 5$ for all simulations and the empirical example.

The sequence γ_k must satisfy the following conditions to ensure convergence of the algorithm:

$$\gamma_k \in (0, 1], \quad \sum_{k=1}^{\infty} \gamma_k = \infty, \quad \text{and} \quad \sum_{k=1}^{\infty} \gamma_k^2 < \infty. \tag{9}$$

These conditions ensure that γ_k decreases slowly to zero, allowing the algorithm to filter out the noise effect and ensuring that the sequence of estimates converges to the root of $\mathbf{g}(\cdot)$. In this study, we set $\gamma_k = \alpha k^{-\beta}$ with some $\alpha > 0$ and $\beta \in (\frac{1}{2}, 1]$, ensuring that the constants decrease slowly to zero.

2.2. Illustrative example

2.2.1. Procedure of bias correction

The dataset generated based on the linear oscillator model is used to illustrate the bias-correction procedure of the two-stage GLLA estimates. For illustrative purposes, we focus on correcting the biased estimate of the frequency parameter η . The RM algorithm takes the following steps to correct the biased estimates:

Step 0: Initialize by setting $k = 0$ and obtaining the initial estimates $\hat{\boldsymbol{\theta}}_0$ using GLLA. Set $\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_0$.

Step 1: Set $k = k + 1$. Generate $\mathbf{U}_1, \dots, \mathbf{U}_m$ from \mathbb{P} and compute:

$$\hat{\mathbf{g}}_m(\boldsymbol{\theta}_k) = \frac{1}{m} \sum_{j=1}^m \hat{\boldsymbol{\theta}}[\mathbf{A}(\boldsymbol{\theta}_k, \mathbf{U}_j)] - \hat{\boldsymbol{\theta}}(\mathbf{y}). \tag{10}$$

⁶This statement is only true when the standard deviation of measurement error σ_e is known, which happens to be the case in the one-parameter illustrative example. However, in general, when σ_e needs to be estimated, it becomes part of the parameter vector $\boldsymbol{\theta}$. In such cases, $e(t)$ should be decomposed as $e(t) = \sigma_e u(t)$, where $u(t) \sim N(0, 1)$, and $u(t)$ is treated as the parameter-free random components.

When generating datasets, the initial values of the displacement (x_1) and the first derivative (\dot{x}_1) are fixed to their values used in the previous simulation illustration for simplicity.
 Step 2: Update θ_k using Equation (7), with $\alpha = 0.3, \beta = 0.6$ for all simulations in this study.
 Step 3: Repeat Steps 1 and 2 until convergence.

The final point estimates after a sufficiently large number of iterations K are obtained using the weighted average

$$\bar{\theta}_K = \frac{\sum_{k=\lfloor K/2 \rfloor + 1}^K \gamma_k \theta_k}{\sum_{k=\lfloor K/2 \rfloor + 1}^K \gamma_k}, \tag{11}$$

where $\lfloor \cdot \rfloor$ denotes the integer part of a real number. In this study, we run a fixed total of 5,000 iterations. Convergence was considered achieved if the relative difference between the point estimates at iterations 4,999 and 5,000 (i.e., their difference divided by the estimate at 4,999) fell below a tolerance threshold of 0.01%.

Additionally, the SE of the frequency parameter was estimated using the following equation (the derivation of a more general form is provided in the Appendix as Equation (16); Equation (12) is a special case of Equation (16)):

$$\widehat{SE} = \sqrt{[\widehat{L}(\bar{\theta}_K)]^{-2} \widehat{\text{Var}}_{\bar{\theta}_K} \{\hat{\theta}\}}, \tag{12}$$

where $\widehat{\text{Var}}_{\bar{\theta}_K} \{\hat{\theta}\}$ was estimated using the MC approximation with 10,000 simulations. The expected value $\mathbb{E}_{\theta} \{\hat{\theta}\}$ is denoted as $L(\theta)$, and $\widehat{L}(\bar{\theta}_K)$ is its derivative evaluated at the final point estimate $\bar{\theta}_K$. This derivative was computed numerically using the finite differences method with the `jacobian` function in the R package `numDeriv` (Gilbert & Varadhan, 2016), where we fixed the parameter-free random components U to ensure numerical variability arises solely from parameter perturbations. This approach is also applied in the computation of the more general form of SE described later. For notational simplicity, from Equation (12) onward, we omit explicit dependence notation in estimators.

2.2.2. Results

Figure 4 displays the iterates of η using the bias-correction method. Although starting from the highly biased GLLA estimate ($\hat{\eta} = -0.618$), the iterates gradually converged, with the final estimate ($\bar{\eta}_K = -0.800$) after $K = 1,000$ iterations matching the true value to the third decimal place. This result demonstrates the effectiveness of the bias-correction procedure in reducing bias in the two-stage GLLA estimates in differential equation models. In addition, the SE estimate for η was 0.169×10^{-3} , indicating high precision in the final estimate.

2.2.3. Summary

This example illustrates the bias-correction procedure in a relatively simple setting, where (a) only one parameter is of focus (i.e., the frequency parameter η in the linear oscillator model) and (b) the model involves only time-independent noise (i.e., measurement error). To assess the performance of the bias-correction method in more general settings, the remainder of this article will extend the approach to the damped linear oscillator model, (a) addressing bias correction and SE estimation for multiple parameters and (b) incorporating time-dependent process noise (i.e., dynamic error) into the model.

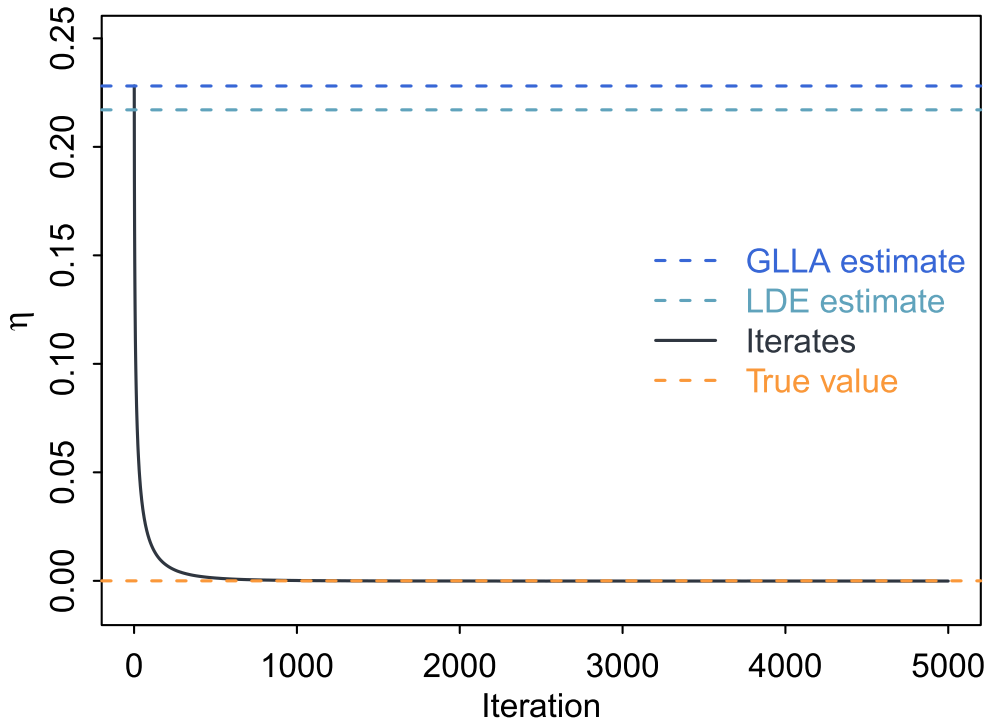


Figure 4. Iterates and point estimates of η in the linear oscillator model.

Note: The solid line represents the iterates of η using the bias-correction method. The dashed lines indicate the true value (orange), the point estimates using GLLA (blue), and the point estimate using LDE (teal).

3. Extension to the damped linear oscillator model

The damped linear oscillator model is a widely used differential equation model in psychological research (e.g., Chow et al., 2005; Ollero et al., 2025; Hu & Treinen, 2019). In this section, we first demonstrate how the bias-correction procedure can be applied to reduce bias in the estimation of multiple parameters in the damped linear oscillator model. We then further evaluate the effectiveness of the bias-correction method by incorporating dynamic error into the model.

3.1. Bias correction for multiple parameters in the damped linear oscillator model

The damped linear oscillator model with measurement error is specified as follows:

$$X(t) = x(t) + e(t), \quad (13)$$

$$\ddot{x}(t) = \eta x(t) + \zeta \dot{x}(t). \quad (14)$$

Equation (13) is the same as Equation (1). In Equation (14), $\dot{x}(t)$ and $\ddot{x}(t)$ represent the first and second derivatives of $x(t)$ with respect to time, respectively. The frequency parameter η describes the frequency of oscillation. The damping parameter ζ governs the amplitude of oscillation.⁷ A negative value of ζ indicates a decrease in oscillation amplitude over time, while a positive value of ζ indicates an increase in amplitude. When $\zeta = 0$, the oscillation amplitude remains constant, and the model reduces to the linear oscillator model.

⁷Although the parameter ζ may be more accurately referred to as a damping parameter when $\zeta < 0$, and an amplification parameter when $\zeta > 0$, we adopt the term “damping parameter” for both signs of ζ , in keeping with the terminology used in most previous research.

3.1.1. Data generation

To preliminarily assess the effectiveness of the bias-correction method for GLLA estimates of multiple parameters, we generated a large dataset based on the damped linear oscillator model, in which the frequency parameter, damping parameter, standard deviation of measurement error, and initial values were all treated as unknown parameters to be estimated. Specifically, individual time-series data ($T = 100$) were generated according to Equations (13) and (14) for $N = 1,000$ individuals. The measurement error-free data were first simulated using the analytic solution of Equation (14), which is valid when $\eta < 0$ and $\zeta^2 + 4\eta < 0$ (Hu & Treinen, 2019):

$$x(t) = e^{\zeta t/2} (A \cos(\omega' t) + B \sin(\omega' t)), \tag{15}$$

where $\omega' = \sqrt{-\eta - \zeta^2/4}$, $A = x_1$, and $B = \dot{x}_1/\omega'$, with x_1 and \dot{x}_1 representing the initial values of the displacement and the first derivative, respectively. We set $\eta = -0.8$ and $\zeta = -0.04$ to reflect a typical cycle length of 7 time points (i.e., $2\pi/\omega' = 2\pi/\sqrt{-\eta - \zeta^2/4} \approx 7$) and a commonly small value of ζ (Cho *et al.*, 2024).

When estimating derivatives using a five-dimensional time-delay embedded matrix via GLLA (see Figure 1), the first row of the resulting matrix represents the derivative estimates at the third time point of the original time series. Therefore, we set the values of the third time point for the zero- and first-order derivatives ($x_3 = -10$ and $\dot{x}_3 = -10$) in order to approximate the time series starting near a positive peak of the oscillation while simultaneously obtaining the true initial values for GLLA. We then used these values as initial conditions to generate $(T - 2)$ time points forward and 2 time points backward according to Equation (15), resulting in T time points for each individual. The observed dataset for each individual was generated by adding measurement error drawn from a normal distribution with mean zero and the standard deviation of $\sigma_e = 1$. This generating process was repeated N times, yielding a final dataset with N individuals, each having T observations.

3.1.2. Estimation and bias correction

To obtain the GLLA estimates of parameters in the damped linear oscillator model, we first estimated the zero-, first-, and second-order derivatives using a five-dimensional time-delay embedded matrix. We then fitted a linear regression model with the zero- and first-order derivatives as predictors and the second-order derivative as the outcome variable to estimate the frequency parameter η and the damping parameter ζ . To estimate the third time point for the zero-order (x_3) and first-order (\dot{x}_3) derivatives, we averaged the first row of the corresponding derivative estimates across individuals. For the standard deviation of measurement error, we used the between-person standard deviation of the first row of the zero-order derivative as an estimate.

To obtain the bias-corrected estimates, we followed the same procedure as in the linear oscillator model, except that in Step 1, five parameters (η , ζ , x_3 , \dot{x}_3 , and σ_e) were estimated and used to generate datasets during each iteration. The final point estimate for each parameter was calculated using Equation (11). The SE estimates were computed as

$$\widehat{SE} = \sqrt{Diag(\widehat{\Sigma})}, \tag{16}$$

where the covariance matrix $\widehat{\Sigma}$ is approximated by (the derivation is provided in the Appendix)

$$\widehat{\Sigma} = \left[\widehat{L}(\bar{\theta}_K) \right]^{-1} \widehat{Cov}_{\bar{\theta}_K} \{ \widehat{\theta} \} \left[\widehat{L}(\bar{\theta}_K) \right]^{-T}. \tag{17}$$

The covariance matrix $\widehat{Cov}_{\bar{\theta}_K} \{ \widehat{\theta} \}$ was estimated using MC approximation with 10,000 simulations, and the Jacobian matrix $\widehat{L}(\bar{\theta}_K)$ was computed using the finite differences method.

For comparison, we also estimated the model parameters using state-space modeling (SSM) implemented in the R package *dynr* (Ou *et al.*, 2019). The initial values for the frequency parameter η and damping parameter ζ were set to the GLLA estimates, while the initial value for the measurement error

Table 1. Parameter estimates in the damped linear oscillator model with and without dynamic error using GLLA, state-space modeling (SSM), and the bias-correction method

Para.	True	Without dynamic error				With dynamic error			
		GLLA	SSM	Bias correction		GLLA	SSM	Bias correction	
		Point Est.	Point Est.	Point Est.	SE Est.	Point Est.	Point Est.	Point Est.	SE Est.
η	-0.8	-0.617	-0.807	-0.799	0.000 ^a	-0.554	-0.763	-0.838	0.029
ζ	-0.04	-0.037	-0.034	-0.040	0.000 ^a	-0.030	-0.111	-0.041	0.007
x_3	-10	-9.461	—	-9.981	0.022	-9.659	—	-10.175	0.048
\dot{x}_3	-10	-6.130	—	-9.981	0.016	-6.229	—	-10.121	0.236
σ_e	1	0.676	0.994	0.971	0.022	0.694	0.623	0.998	0.011
σ_ϵ	0.5	—	—	—	—	1.274	5.663	0.302	0.167

Note: ^a The values of 0.000 are a result of rounding to three decimal places. When rounded to four decimal places, the values are 0.0001.

standard deviation σ_e was set to the between-person standard deviation of the first row of the zero-order derivative estimates, consistent with the approach used for GLLA and the bias-correction method. For the initial conditions, which in *dynr* correspond to the first time point of the latent states (i.e., the zero- and first-order derivatives), we used the sample means across individuals of the first time point values for the zero-order derivative, and the mean of the numerical derivatives between the first and second time points for the first-order derivative. Given that the SSM estimates initial conditions at the first time point, whereas the bias-correction method estimates them at the third time point, a direct comparison of these initial condition parameters is not feasible. We imposed appropriate boundary constraints, restricting $\sigma_e > 0$ and $\eta < 0$. All other estimation settings were kept at their default values.

3.1.3. Results

Figure 5 shows the iterates of parameters using the bias-correction method. All parameter iterates converged to equilibrium within 5,000 iterations. Compared to the GLLA estimates, the biased-corrected estimates were much closer to the true values, and exhibited small SE estimates (see Table 1). The SSM estimates for the frequency parameter η and the measurement error standard deviation σ_e were also highly accurate. However, for the damping parameter ζ , the SSM estimate exhibited substantial bias, which was even larger than that of the GLLA estimate. This suggests that the bias-correction method effectively reduces the bias of the GLLA estimates for multiple parameters in the damped linear oscillator model with only time-independent noise (i.e., measurement error). This result motivates further investigation into the performance of the bias-correction method when time-dependent dynamic error is introduced into the model.

3.2. Bias correction in the damped linear oscillator model with dynamic error

The damped linear oscillator model with both time-independent measurement error and time-dependent dynamic error can be expressed as follows:

$$X(t) = x(t) + e(t), \quad (18)$$

$$\ddot{x}(t)\Delta t = [\eta x(t) + \zeta \dot{x}(t)]\Delta t + \Delta \epsilon(t). \quad (19)$$

Equation (18) is identical to Equations (1) and (13). In Equation (19), the original damped linear oscillator model in Equation (14) is multiplied by Δt with the addition of the change in dynamic error

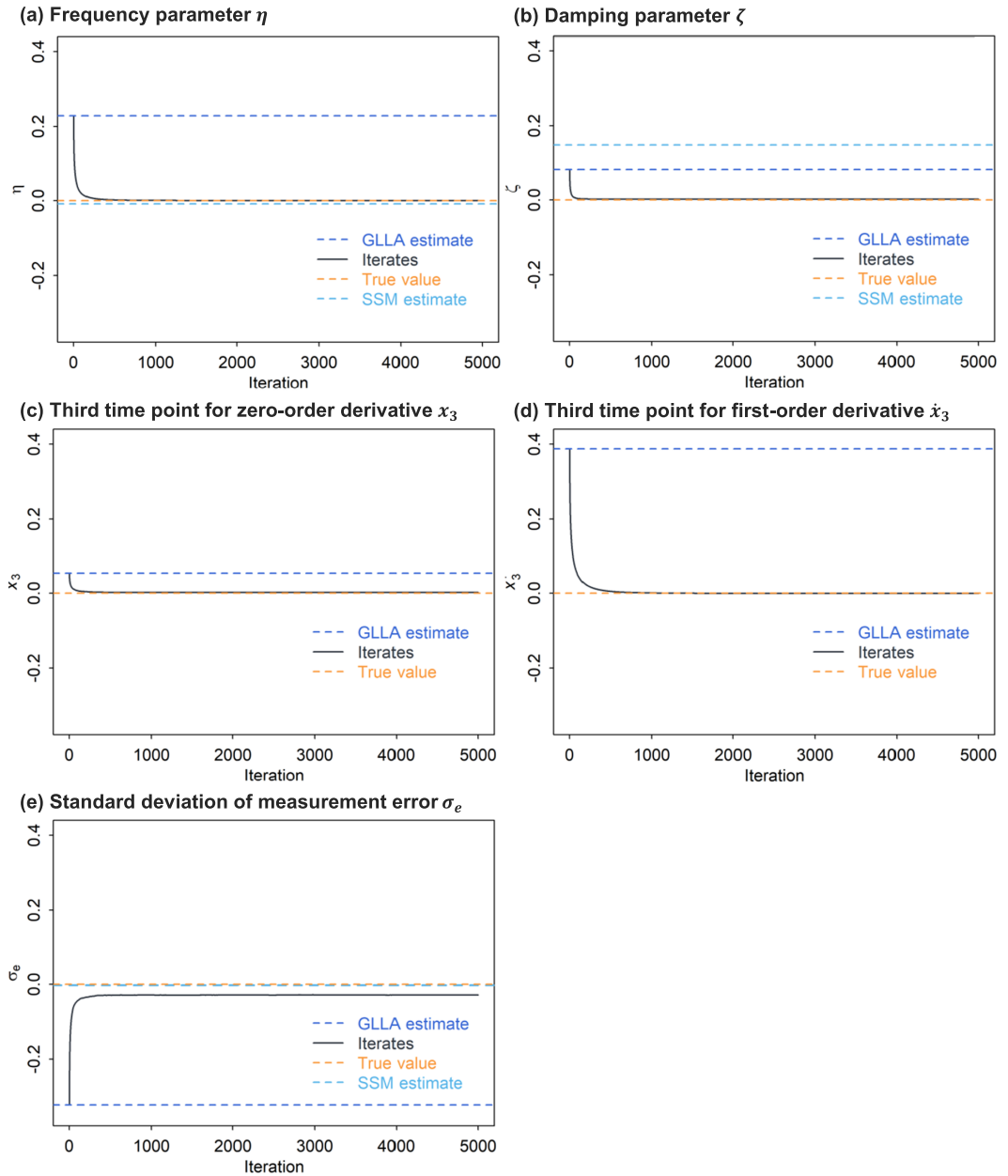


Figure 5. Relative bias of the parameter estimates and iterates in the damped linear oscillator model.
Note: The solid line represents the iterates of parameters using the bias-correction method. The dashed lines indicate the true value (orange), the GLLA estimates (dark blue), and the SSM estimates (light blue).

$\Delta\varepsilon(t)$, resulting in a stochastic version of the model.⁸ The change in dynamic error $\Delta\varepsilon(t)$ over a short time interval $[t, t + \Delta t]$ follows a normal distribution with mean zero and variance $\sigma_e^2 \Delta t$. In other words, $\Delta\varepsilon(t)$ can be modeled as the increment of a univariate standard Wiener process $W(t)$ scaled by the

⁸Another typical stochastic formulation of the damped oscillator model involves two first-order differential equations:

$$\begin{aligned} \text{footnoted}x(t) &= h(t) dt, \\ dh(t) &= [\eta x(t) + \zeta h(t)] dt + \sigma_e dW(t), \end{aligned}$$

standard deviation σ_ε , such that $\Delta\varepsilon(t) = \sigma_\varepsilon\Delta W(t)$, where $\Delta W(t) = W(t + \Delta t) - W(t)$ follows a normal distribution with mean zero and variance of Δt (i.e., $\Delta W(t) \sim N(0, \Delta t)$).

Although both dynamic error $\varepsilon(t)$ and measurement error $e(t)$ follow normal distributions, they serve fundamentally different roles in the model. The dynamic error captures stochastic shocks that continuously influence the current and future states of the dynamic system, while the measurement error represents noise that only affects the current state (i.e., the measurement occasion) of the system (Chen et al., 2018). In psychological studies that use differential equation models to examine dynamic processes, incorporating dynamic error is a more realistic assumption. This assumption reflects that the current state of the dynamic system cannot be perfectly predicted by its previous state, as random noise can contribute to the dynamic process at any moment (Ollero et al., 2025). Moreover, it accounts for the impact of unknown or unmodeled factors on the system's evolution.

3.2.1. Data generation

To generate data for the damped linear oscillator model with dynamic error, we first simulate the measurement-error-free time-series data for each individual using the Euler method for numerical integration. This method approximates the state of the system over discrete time steps with the following update rules:

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t, \quad (20)$$

$$\dot{x}(t + \Delta t) = \dot{x}(t) + [\eta x(t) + \zeta \dot{x}(t)]\Delta t + \Delta\varepsilon(t), \quad (21)$$

where the time step size is $\Delta t = 0.01$. The initial conditions for the zero- and first-order derivatives were set at the third time point (x_3 and \dot{x}_3). Numerical integration was used to generate $(T - 2)/\Delta t$ time points forward and $2/\Delta t$ time points backward, where $T = 1,000$, resulting in a total of $T/\Delta t = 100,000$ states. A thinning interval of $1/\Delta t$ was then applied, resulting in $T = 1,000$ data points for each individual. The observed dataset for each individual was generated by adding measurement error drawn from a normal distribution with mean zero and standard deviation σ_e . This process was repeated $N = 100$ times, yielding a final dataset with N individuals, each having T observations. The true values of parameters were set to the same values used in the previous simulation ($\eta = -0.8$, $\zeta = -0.04$, $x_3 = -10$, $\dot{x}_3 = -10$, and $\sigma_e = 1$), and the standard deviation of the dynamic error was set to $\sigma_\varepsilon = 0.5$.

3.2.2. Estimation and bias correction

The GLLA estimates of the parameters η , ζ , x_3 , \dot{x}_3 , and σ_e were obtained using the same method as in the previous section. Additionally, the standard deviation of dynamic error σ_ε was estimated using the residual standard deviation from the linear regression of the zero- and first-order derivatives on the second-order derivative. The bias-corrected estimates were obtained for all six parameters (η , ζ , x_3 , \dot{x}_3 , σ_e , and σ_ε). The point estimates and SE estimates computed according to Equations (11) and (16), respectively.

We also estimated the model parameters using SSM method. We specified the model following the example provided in the package documentation (<https://github.com/mhunter1/dynr/blob/master/demo/LinearSDE.R>). The initial values for the five parameters (η , ζ , x_3 , \dot{x}_3 , and σ_e) were set following the same procedure as described in the previous section for the model without dynamic error. For the dynamic error standard deviation σ_ε , the initial value was set to the GLLA estimate, consistent with

where $h(t)$ is the first derivative of $x(t)$ with respect to time, and $dW(t)$ is the "differential" of a standard univariate Wiener process that is defined via Itô's calculus. However, to align with the second-order differential form used throughout this article, we express the stochastic model by multiplying the deterministic equation by Δt and adding a dynamic error term $\Delta\varepsilon(t)$, which captures the change in dynamic error over the interval $[t, t + \Delta t]$.

the approach used for the bias-correction method. We imposed boundary constraints, requiring $\sigma_e > 0$, $\sigma_\varepsilon > 0$, and $\eta < 0$.

3.2.3. Results

Figure 6 presents the parameter iterates using the bias-correction method in the damped linear oscillator model with dynamic error. All parameters converged to their equilibrium within 5,000 iterations. As shown in Table 1, the GLLA estimates of the frequency parameter η and the damping parameter ζ in the damped linear oscillator model with dynamic error exhibited larger bias than those in the model without dynamic error. The GLLA estimate of σ_ε in the model with dynamic error also showed a large deviation from its true value. The SSM estimates for the frequency parameter η and the measurement error standard deviation σ_e showed acceptable bias levels. However, consistent with the results from the model without dynamic error, the SSM estimate for the damping parameter ζ exhibited substantial bias that was larger than that of the GLLA estimate. Furthermore, the SSM estimate for the dynamic error standard deviation σ_ε showed severe overestimation. For the biased-corrected estimates, although some still deviated slightly from their true values, the bias was largely reduced compared to the initial estimates using GLLA. The SE estimates were relatively larger than those in the damped linear oscillator model without dynamic error (except for the SE estimate of σ_e). These findings provide initial evidence for the effectiveness of the bias-correction method in addressing multiple parameters in a differential equation model with dynamic error.

4. Simulation study

Next, we conducted a simulation study to examine the effectiveness of the bias-correction method in the damped linear oscillator model with dynamic error. Specifically, we varied the true values of the damping parameter, the sample size, and the number of time points per individual, and compared the parameter estimates using GLLA, SSM, and the bias-correction method to assess their performance across different scenarios.

4.1. Procedure

We adopted the same procedure of data generation and parameter estimation as described in Section 3.2. In summary, we first simulated time series for each individual using the Euler method for numerical integration, and added measurement error to obtain observed data. This process was repeated for N individuals, each having T observations. For parameter estimation, we first constructed a five-dimensional time-delay embedded matrix and applied GLLA to estimate zero-, first-, and second-order derivatives. Using the estimated derivatives, we obtained the GLLA estimates of the six parameters (η , ζ , x_3 , \dot{x}_3 , σ_e , and σ_ε) in the damped linear oscillator model with dynamic error. Bias correction was performed using the RM algorithm outlined in Section 2.2.1. The point estimates were computed according to Equation (11), and the SE estimates were obtained using MC approximation and numerical differentiation method according to Equation (16). For comparison, we also estimated the parameters based on state-space models using the *dynr* package (Ou *et al.*, 2019), following the same procedure as described in the previous section. The analysis was conducted in R 4.2.2 (R Core Team, 2020).

4.2. Simulation conditions

We varied the damping parameter ($\zeta = -0.04, 0, 0.04$), the sample size ($N = 50, 150, 500$), and the number of time points per individual ($T = 14, 28, 56$), which were fully crossed, resulting in 27 ($= 3 \times 3 \times 3$) simulation conditions. In each condition, we replicated the MC simulation 500 times. The other parameters ($\eta = -0.8$, $x_3 = -10$, $\dot{x}_3 = -10$, $\sigma_e = 1$, and $\sigma_\varepsilon = 0.5$) were held constant across all simulation conditions. The frequency parameter η and the damping parameter ζ are of primary interest in the

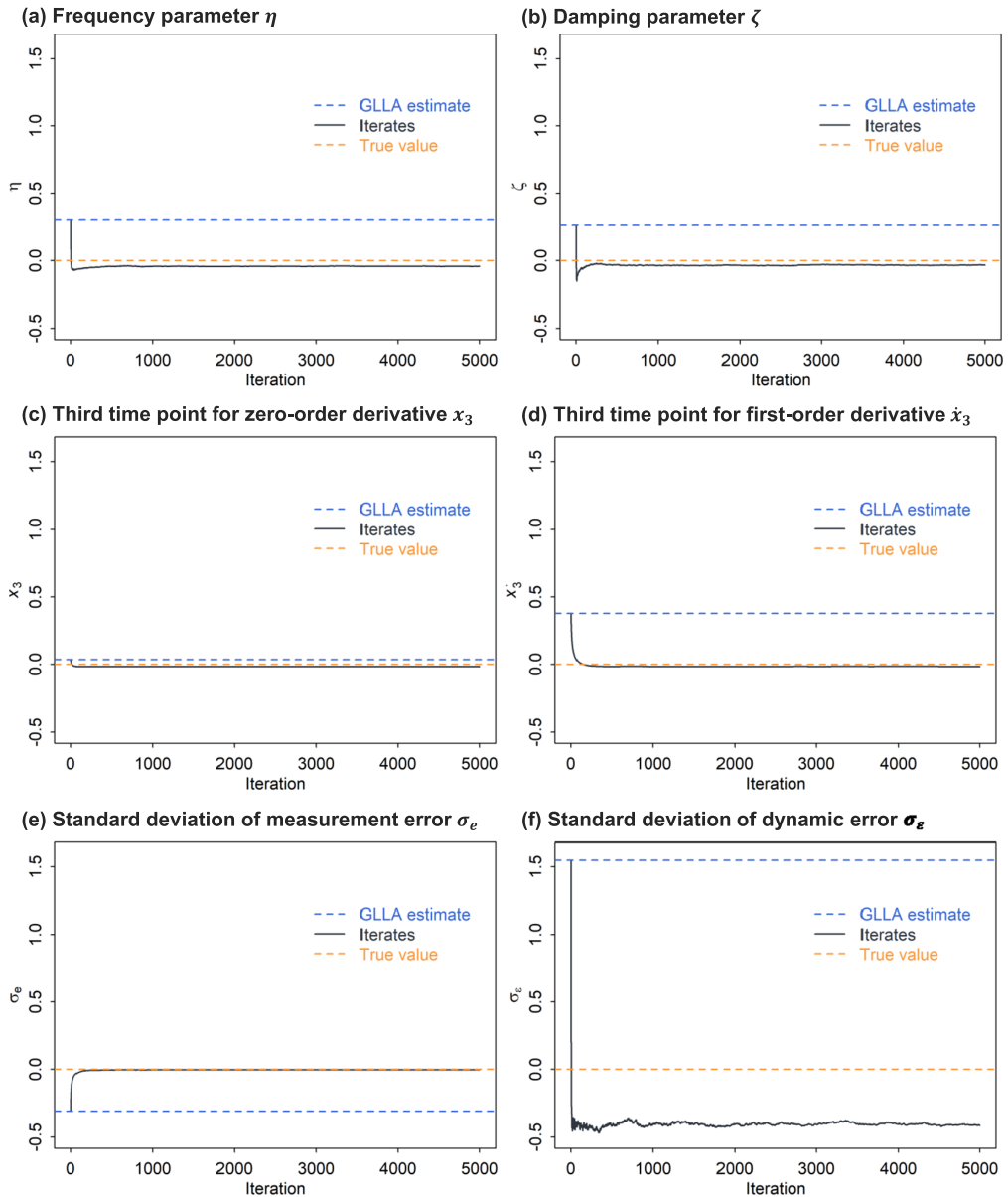


Figure 6. Relative bias of the parameter estimates and iterates in the damped linear oscillator model with dynamic error.

Note: The solid line represents the iterates of parameters using the bias-correction method. The dashed lines indicate the true value (orange), the GLLA estimates (blue), and the SSM estimates (purple). The SSM point estimates (see Table 1) are not included in this figure to maintain clarity in visualization, as the large relative bias for σ_e would disproportionately scale the y-axis.

damped linear oscillator model. In the simulation study, we focused on $\eta = -0.8$ to account for a typical cycle length of 7 time points (e.g., one week with one measurement per day), and varied ζ to represent a system with a damping process ($\zeta < 0$), an amplification process ($\zeta > 0$), or without damping ($\zeta = 0$). For sample size, we used 50, 150, and 500 to represent small, medium, and large samples, respectively, as in typical intensive longitudinal studies. For the number of time points per individual, we chose 14, 28, and 56 to represent common study durations, covering approximately two, four (about one month), and eight (about two months) cycles.

4.3. Evaluation criteria

4.3.1. Convergence rate

For the bias-correction method, we run a fixed total of 5,000 iterations. Convergence was determined by computing the relative difference between the point estimates at iterations 4,999 and 5,000 (i.e., their difference divided by the estimate at 4,999), with a tolerance threshold set to 0.01%.⁹ For the converged cases of the bias-correction and SSM methods, we further examined the SE estimates to exclude outliers or improper solutions (i.e., SE estimates greater than 10 were considered outliers, as they were typically much less than 10). For the bias-correction method, we identified one and two cases containing outliers in the following two simulation conditions, respectively: (a) $N = 500$, $T = 56$, $\zeta = 0.04$ and (b) $N = 50$, $T = 56$, $\zeta = 0$. For the SSM method, we identified one and two cases containing outliers in the following two simulation conditions, respectively: (a) $N = 150$, $T = 14$, $\zeta = 0$ and (b) $N = 150$, $T = 14$, $\zeta = 0.04$. The convergence rate for each method and condition was calculated as the percentage of converged replications without outliers across the 500 simulation replications.

4.3.2. Point estimates

For the six parameters (η , ζ , x_3 , \dot{x}_3 , σ_e , and σ_ε), we compare the accuracy of estimates using GLLA and the bias-correction method. For the parameters η , ζ , σ_e , and σ_ε , we also report the estimation results from the SSM method. Specifically, we calculated the bias (i.e., the deviation of the average estimates across all replications from the true parameter values) for parameters with true values of 0 (e.g., $\zeta = 0$ in some simulation conditions), and the relative bias (i.e., bias divided by the true values) for parameters with non-zero true values. The MC SEs of the bias and relative bias were computed based on the converged replications without outliers. Additionally, we computed the root mean squared error (RMSE) of the parameter estimates.

4.3.3. Standard error estimates

We used the standard deviation of the parameter estimates from the converged replications without outliers as the true values of SE estimates (i.e., empirical SEs). For the bias-correction method, the bias, relative bias, and RMSE of the SE estimates were calculated to assess the accuracy of the SE estimates.

4.4. Results

Both the bias-correction and SSM methods showed high convergence rates across all conditions, ranging from 96.6% to 100% (Mean (M) = 99.33%). Table 2 presents the parameter estimation results in three example conditions ($N = 150$, $T = 28$, $\zeta = -0.04, 0, 0.04$). The complete results of the simulation study are provided in Section S4 of the Supplementary Material. The key findings are summarized below.

Figure 7 presents the relative bias of parameters estimated using GLLA, SSM, and the bias-correction method. The results indicated that GLLA consistently underestimated the frequency parameter η , the third time points of the zero-order and first-order derivatives (x_3 and \dot{x}_3), and the standard deviation of measurement error σ_e , while overestimated the standard deviation of dynamic error σ_ε . For the damping parameter ζ , it was underestimated when the true value of $\zeta < 0$, overestimated when $\zeta > 0$, and showed lower bias when $\zeta = 0$.

The SSM method showed acceptable bias for the frequency parameter η , though the bias was somewhat larger when $\zeta = -0.04$ and $T = 56$. For the damping parameter ζ , the SSM estimates were accurate when the true value was 0, but showed varying degrees of bias when $\zeta = 0.04$ or -0.04 , with noticeably larger MC SEs in some conditions. For the measurement error standard deviation σ_e , the SSM estimates were accurate when $T = 14$ but became increasingly biased (tending toward underestimation)

⁹The iterative process of the bias-corrected estimates is presented in Section S3 of the Supplementary Material, which supports the sufficiency of 5,000 iterations for achieving stable estimates under our simulation conditions.

Table 2. Parameter estimates using GLLA, the bias-correction method, and state-space modeling (SSM) in example conditions

Para.	Type	True	GLLA			Bias correction			SSM		
			Est	rBias	RMSE	Est	rBias	RMSE	Est	rBias	RMSE
η	Damping	-0.8	-0.584	-0.271 [-0.275, -0.267]	0.219	-0.810	0.012 [0.009, 0.015]	0.029	-0.776	-0.030 [-0.035, -0.026]	0.049
	Constant	-0.8	-0.595	-0.256 [-0.259, -0.253]	0.207	-0.805	0.007 [0.004, 0.009]	0.025	-0.785	-0.019 [-0.022, -0.015]	0.035
	Amplifying	-0.8	-0.602	-0.248 [-0.250, -0.245]	0.199	-0.804	0.005 [0.002, 0.008]	0.026	-0.790	-0.012 [-0.015, -0.009]	0.026
ζ	Damping	-0.04	-0.026	-0.343 [-0.378, -0.309]	0.021	-0.041	0.013 [-0.024, 0.050]	0.017	-0.063	0.584 [0.491, 0.676]	0.047
	Constant	0	0.011	-	0.016	0.000	-	0.013	-0.010	-	0.033
	Amplifying	0.04	0.047	0.175 [0.152, 0.198]	0.012	0.040	0.011 [-0.014, 0.036]	0.011	0.038	-0.046 [-0.096, 0.004]	0.022
x_3	Damping	-10	-9.535	-0.047 [-0.047, -0.046]	0.470	-10.008	0.001 [0.000, 0.002]	0.078	-	-	-
	Constant	-10	-9.577	-0.042 [-0.043, -0.042]	0.428	-10.003	0.000 [-0.000, 0.001]	0.078	-	-	-
	Amplifying	-10	-9.619	-0.038 [-0.039, -0.037]	0.386	-10.000	-0.000 [-0.001, 0.001]	0.076	-	-	-
\dot{x}_3	Damping	-10	-6.267	-0.373 [-0.375, -0.372]	3.736	-10.069	0.007 [0.005, 0.009]	0.256	-	-	-
	Constant	-10	-6.126	-0.387 [-0.389, -0.386]	3.877	-10.039	0.004 [0.002, 0.006]	0.269	-	-	-
	Amplifying	-10	-6.007	-0.399 [-0.401, -0.398]	3.996	-10.039	0.004 [0.002, 0.006]	0.249	-	-	-
σ_e	Damping	1	0.696	-0.304 [-0.308, -0.300]	0.307	1.000	-0.000 [-0.003, 0.002]	0.030	0.860	-0.140 [-0.164, -0.117]	0.298
	Constant	1	0.697	-0.303 [-0.307, -0.300]	0.306	1.000	0.000 [-0.002, 0.003]	0.029	0.880	-0.120 [-0.145, -0.096]	0.297
	Amplifying	1	0.696	-0.304 [-0.308, -0.300]	0.307	1.000	-0.000 [-0.003, 0.002]	0.030	0.854	-0.146 [-0.172, -0.120]	0.327
σ_ϵ	Damping	0.5	1.164	1.329 [1.235, 1.423]	0.846	0.535	0.070 [0.013, 0.127]	0.320	1.905	2.809 [2.603, 3.015]	1.814
	Constant	0.5	1.137	1.273 [1.181, 1.366]	0.819	0.505	0.011 [-0.042, 0.063]	0.292	1.808	2.616 [2.413, 2.818]	1.730
	Amplifying	0.5	1.185	1.370 [1.270, 1.470]	0.884	0.522	0.044 [-0.010, 0.098]	0.301	1.887	2.773 [2.563, 2.983]	1.815

Note: Type corresponds to the true value of ζ : Damping ($\zeta = -0.04$), Constant ($\zeta = 0$), and Amplifying ($\zeta = 0.04$). Est denotes the mean of estimates across all converged replications without outliers. rBias denote the bias and relative bias (with ± 2 Monte Carlo standard error intervals), respectively. RMSE denotes the root mean square error.

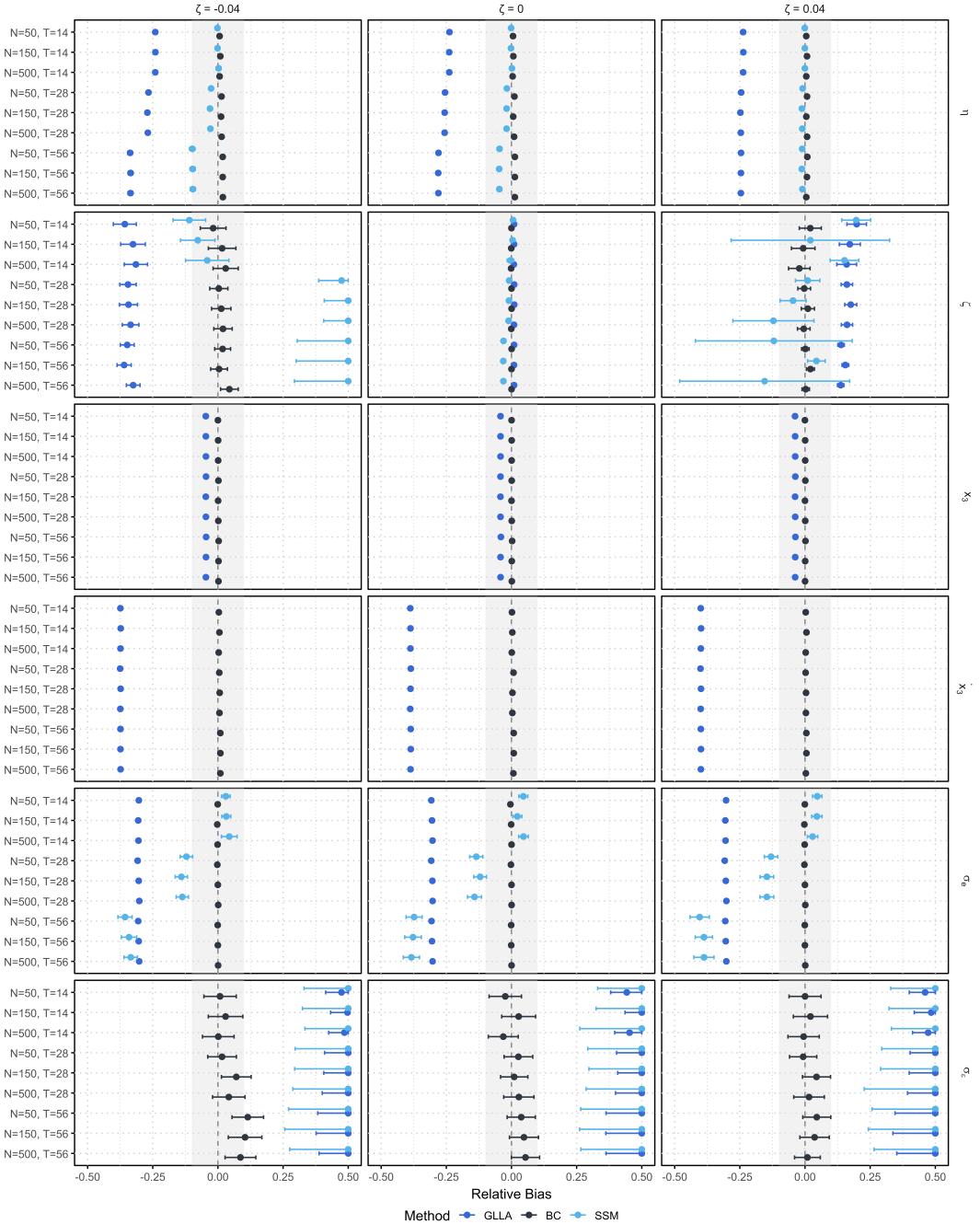


Figure 7. Relative bias of parameters estimated using GLLA, the bias-correction (BC) method, and the state-space modeling (SSM) method.

Note: The error bars show the ± 2 Monte Carlo standard error intervals. Values exceeding 0.5 are shown as 0.5 for better visualization. The gray area indicates the acceptable range of relative bias ($\pm 10\%$). When the true value of ζ is 0, the result for ζ is reported as bias rather than relative bias.

as T increased. Consistent with the demonstrations in the damped linear oscillator model, the SSM estimates for the dynamic error standard deviation σ_ε showed substantial overestimation across all conditions.

In contrast, the bias-corrected estimates for all six parameters showed high estimation accuracy, with the absolute values of relative bias lower than 5% for η , ζ , x_3 , and \dot{x}_3 in most conditions, and lower than 10% for σ_ε in all but two specific conditions, where the relative bias was 10.4% and 11.5%.

Figure 8 shows the RMSE of parameters estimated using the three methods. The RMSE for parameters estimated using the bias-correction method was consistently lower than or similar to those of the SSM and GLLA methods. These results support the effectiveness of the bias-correction method in reducing the bias of the GLLA estimates.

For the SE estimates, the bias-correction method showed small empirical SEs for η (ranging from 0.020 to 0.048, $M = 0.030$), ζ (ranging from 0.006 to 0.023, $M = 0.015$), x_3 (ranging from 0.059 to 0.126, $M = 0.089$), and σ_ε (ranging from 0.015 to 0.052, $M = 0.032$), while relatively higher empirical SEs were observed for \dot{x}_3 (ranging from 0.246 to 0.304, $M = 0.280$), and σ_ε (ranging from 0.276 to 0.366, $M = 0.324$). The absolute bias of SE estimates was less than 0.010 for η , ζ , x_3 , and σ_ε , and less than 0.025 for \dot{x}_3 and σ_ε .

Overall, the bias of the GLLA estimates was largely reduced after applying the bias-correction method. Additionally, the SE estimates using the bias-correction method demonstrated relatively high precision in parameter estimates. These findings suggest that the bias-correction method can effectively estimate parameters in the damped linear oscillator model with dynamic error.

5. Empirical illustration

5.1. Data

The dataset comprises daily sales records from 119 stores in a shopping mall, tracked over a period of 14 days following a promotional event. This data was previously utilized by Hu and Treinen (2019) to illustrate the damped oscillatory patterns in consumer behavior following sales promotions.

5.2. Analysis

We first standardized the sales for each individual store using the mean and standard deviation of each store's sales over the 14-day period (i.e., within-store standardization). Next, we detrended the standardized sales by fitting a simple linear regression against time (with the promotion day as $t = 1$) for each store, and saved the residuals for further analysis. We then fitted a damped linear oscillator model with dynamic error. Parameters were estimated using GLLA and the proposed bias-correction method. The bias-correction procedure followed the same settings as in the simulation study, except that the number of iterations in the RM algorithm was fixed at 10,000 to obtain more stable results. Finally, we generated 10,000 simulated trajectories based on both the GLLA estimates and the bias-corrected estimates, plotting the median trajectories (as fitted curves) and their 95% prediction intervals.

5.3. Results

Parameter estimation results revealed substantive differences between methods (see Table 3). The bias-correction method yielded estimates of $\eta = -0.864$ for the frequency parameter and $\zeta = -0.472$ for the damping parameter, compared to $\eta = -0.526$ and $\zeta = -0.266$ from GLLA. While both methods indicate a damped oscillatory process of daily sales after the promotion day, the fitted trajectories differ substantially. As shown in Figure 9, the bias-corrected median trajectory exhibited quicker oscillatory decay compared to the GLLA-based trajectory.

In addition, clear differences were observed in the prediction intervals between methods. The 95% prediction interval based on the bias-correction method was not only narrower than that based on GLLA, but also showed much less widening over time. In contrast, the prediction interval based

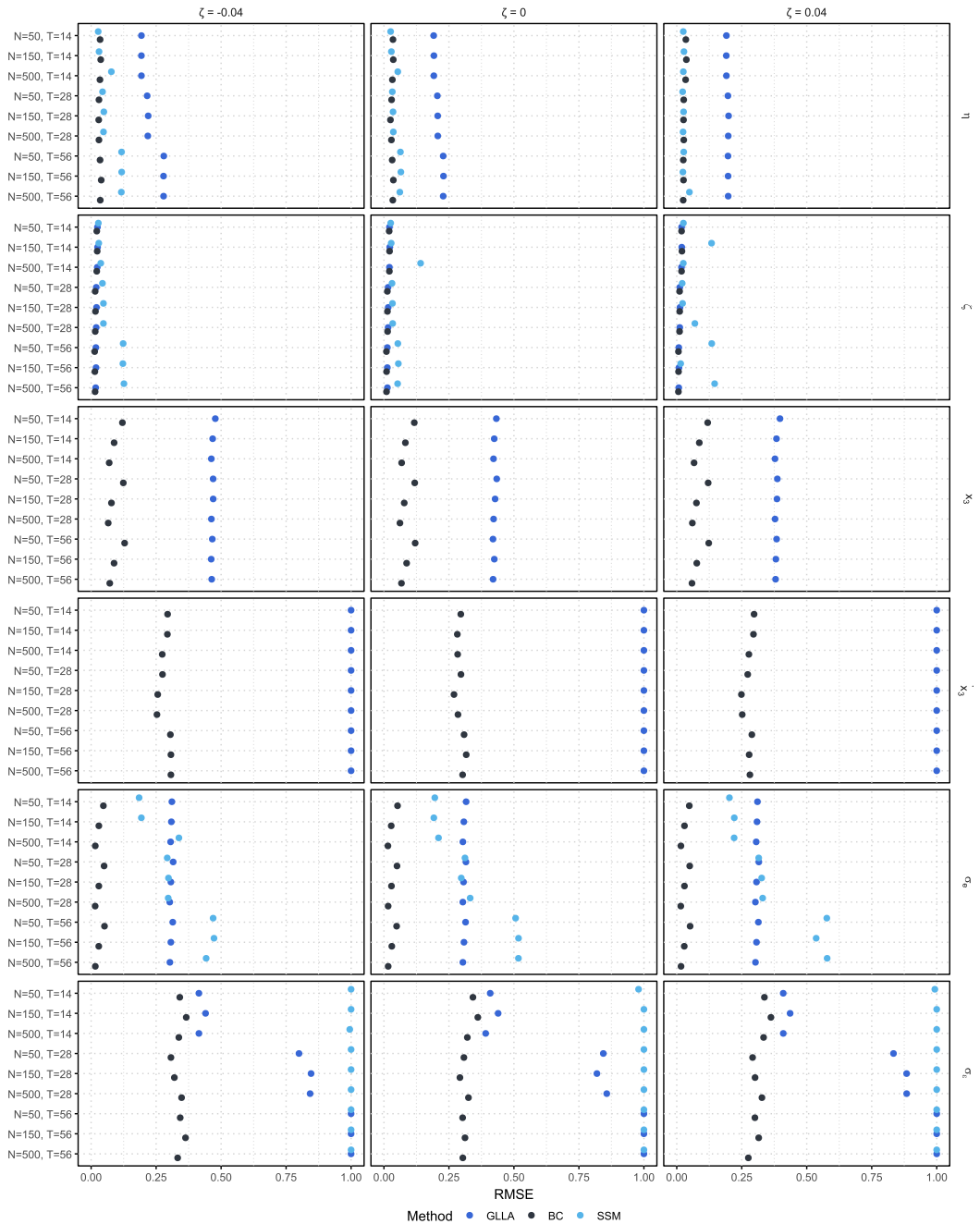


Figure 8. RMSE of parameters estimated using GLLA, the bias-correction (BC) method, and the state-space modeling (SSM) method. *Note:* Values exceeding 1 are shown as 1 for better visualization.

on GLLA expanded substantially as time progressed. These differences reflect the discrepancy in the estimated standard deviation of dynamic error σ_ϵ : GLLA produced a considerably larger estimate of $\sigma_\epsilon = 0.219$, compared to $\sigma_\epsilon = 0.107$ from the bias-correction method. Because dynamic error continuously accumulates uncertainty over time, a larger σ_ϵ amplifies the growth of the prediction interval, especially at later time points. It is important to note, however, that a smaller estimate of σ_ϵ does not necessarily

Table 3. Parameter estimates in the damped linear oscillator model with dynamic error for sales data using GLLA and the bias-correction method

Parameter	GLLA		Bias correction	
	Estimate	Estimate	Estimate	SE
η	-0.526	-0.864	-0.864	0.159
ζ	-0.266	-0.472	-0.472	0.125
x_3	-0.363	-0.417	-0.417	0.038
\dot{x}_3	-0.591	-0.791	-0.791	0.126
σ_e	0.355	0.715	0.715	0.023
σ_ε	0.219	0.107	0.107	0.078

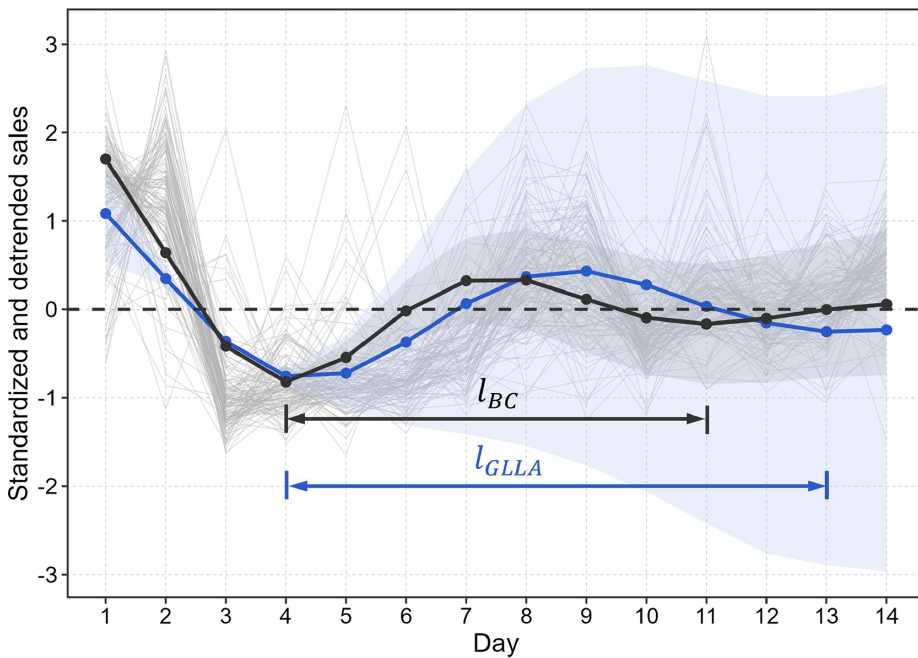


Figure 9. Comparison of fitted trajectories for standardized and detrended sales based on the GLLA estimates and the bias-corrected estimates.

Note: Light gray lines depict the observed daily sales trajectories for all 119 stores. Solid black line represents the median trajectory based on bias-corrected estimates (gray area indicates 95% prediction interval). Solid blue line shows the GLLA-based median trajectory (blue area indicates 95% prediction interval). The temporal distance (in days) between the two lowest points of the median trajectories are highlighted.

imply a more accurate representation of the underlying process, as the true value is unknown in the empirical data. Our purpose here is not to claim superiority of one method over the other based on dynamic error variance, but to illustrate how differences in parameter estimates are manifested in the fitted trajectories and their prediction intervals.

To further characterize the sales dynamics, we approximated the cycle length of sales dynamics using the formula $2\pi/\sqrt{-\eta - \zeta^2/4}$ (i.e., derived without dynamic error). The cycle length computed using the GLLA estimates was 8.563 days, while the bias-corrected estimates yielded a cycle length of 6.987 days. This difference is also evident in the fitted trajectories (see Figure 9), where the temporal distance between the two lowest points was seven days (l_{BC}) for the bias-corrected median trajectory and nine

days (l_{GLLA}) for the GLLA-based trajectory. Given that shopping behavior typically follows weekly cycles with peaks on weekends, the cycle length approximated using the bias-corrected estimates aligns more closely with the common weekly shopping pattern. These results suggest that the bias-correction method provides improved accuracy in capturing the underlying dynamics of post-promotion sales.

6. Discussion

Differential equation models are powerful tools for investigating dynamic processes. These models provide a framework for understanding the relations between variables and their derivatives with respect to time, offering insights into the underlying mechanisms such as self-regulation. Among the various estimation methods, the GLLA is a widely used two-stage method due to its simplicity and ease of implementation. However, GLLA often produces biased parameter estimates. This study proposed a bias-correction method for GLLA via SA, which produces asymptotically unbiased estimates even when the initial bias of GLLA estimates is substantial. We also provide valid asymptotic SEs for the bias-corrected estimates.

6.1. Summary of findings

Through a simulation illustration, we first showed that GLLA, while straightforward, yielded biased estimates of a key parameter (i.e., the frequency parameter η) in the linear oscillator model. Specifically, we found that the bias in GLLA estimates of η becomes more pronounced as the absolute true value of η increases (see Figure 3). This is particularly concerning because, in real-world applications, researchers often encounter dynamic systems with high-frequency oscillations, where η can take on large absolute values (e.g., with η estimated to be between -0.4 and -0.9 ; Cho *et al.*, 2024). To address this issue, we applied the proposed bias-correction method, which successfully reduced the bias in the GLLA estimate of η , yielding highly accurate and precise results. This initial demonstration highlights the limitations of GLLA and suggest the potential of the bias-correction approach to more complex differential equation models.

Building on this foundation, we extended the bias-correction method to the damped linear oscillator model, introducing additional complexity by simultaneously accounting for multiple parameters and incorporating time-dependent dynamic error (or process noise). In the damped linear oscillator model, the frequency parameter (η) and the damping parameters (ζ) are often of primary interest, but other parameters, such as the initial conditions (x_3 and \dot{x}_3) and the standard deviation of measurement error (σ_ϵ) and dynamic error (σ_ϵ), also important in the estimation process, and have substantial meanings that may be of interests. Our simulation study demonstrated that the bias-correction method could reduce bias in the GLLA estimates of all these parameters, offering an effective approach to improving parameter estimation in the damped linear oscillator model.

In addition to addressing bias in the estimation of multiple parameters, we further extended the bias-correction method to incorporate dynamic error, accounting for the stochastic nature of real-world processes. Dynamic error represents random innovations that continuously influence the system's state, capturing the inherent unpredictability of many psychological processes (Chen *et al.*, 2018; Ollero *et al.*, 2025). Our simulation results showed that the proposed bias-correction method was able to handle the additional complexity introduced by dynamic error, demonstrating its ability to provide accurate parameter estimates even in the presence of stochastic variability. By accounting for this source of variability, our method offers a more realistic representation of dynamic processes in differential equation models.

To evaluate the robustness of the bias-correction method in the damped linear oscillator model with dynamic error, we conducted a simulation study across a range of conditions, including different sample sizes, numbers of time points, and values of the damping parameter (ζ). The results showed that the bias-correction method consistently outperformed GLLA, with (much) smaller bias in parameter estimation. The bias-correction method also showed relatively high estimation precision, with small

bias of SE estimates. Notably, the method performed well even with a small number of time points (e.g., 14) and relatively small sample size (i.e., 50), which are not uncommon in intensive longitudinal studies. Furthermore, the method showed consistent performance across different values of ζ , including cases where ζ was negative (damping), positive (amplification), or zero (no damping), indicating its effectiveness under various dynamic mechanisms.

We also compared our bias-correction approach with SSM. SSM represents a powerful and theoretically grounded one-stage framework for estimating differential equation models. However, our findings suggest that in certain practical scenarios, SSM may still encounter estimation accuracy challenges. While SSM provided accurate estimates for the frequency parameter η in many conditions, it showed substantial bias in estimating the damping parameter ζ and the dynamic error standard deviation σ_ϵ . This is consistent with recent findings by Ollero et al. (2025), who highlighted that continuous-time state-space models may struggle with accuracy when the number of time points is relatively small. Our results further demonstrate that the proposed bias-correction method maintains high accuracy even with relatively few time points (e.g., $T = 14$ or 28), underscoring its potential advantage for short time series. Additionally, as other one-stage methods, SSM requires appropriate initial values and sometimes may even need carefully specified boundary constraints (e.g., setting upper bounds on parameter values) beyond simple positivity constraints to achieve satisfactory estimation results. In contrast, our bias-correction method demonstrates relative robustness and high accuracy across various conditions without requiring extensive fine-tuning of boundary conditions.

6.2. Contributions and implications

Taken together, the findings of this study demonstrate the effectiveness of the bias-correction method in reducing bias of two-stage estimates in differential equation models. While the current study focused on second-order differential equation models such as the damped linear oscillator model, the underlying principles of the bias-correction method suggest its potential for broader applications. Specifically, the proposed method requires only an initial parameter estimation procedure (even if biased) and a rule for generating data based on the estimated parameters, enabling iterative bias reduction through SA. This opens up new possibilities for applying the method to a wide range of models. For instance, it could be used in differential equation models for bivariate or multivariate processes. Previous research has employed coupled differential equation models (Hu et al., 2014) to examine the interactions between two individuals (e.g., romantic couples or parent–child dyads), or the predator–prey model (Lotka, 1920) to explore negative feedback mechanisms in dynamic bidirectional relations. The two-stage GLLA has also been used to estimate parameters in these models in psychological research (Luo & Hu, 2023). The procedure of bias correction for the GLLA estimates in these models follows the same principles as in the current study, despite the increased number of parameters to be estimated. Given the flexibility and adaptability of the proposed bias-correction method, future research could refine the method and explore its applications to more complex models, enhancing our understanding of dynamic behavioral mechanisms.¹⁰

In addition, our findings highlight the need for caution when applying GLLA, as well as other two-stage estimation methods, in differential equation models. As shown in our initial illustration, systematic bias was observed in the frequency parameter estimated using GLLA. It is important to note that the accuracy of such two-stage methods is influenced by the oscillation frequency relative to the sampling rate. Specifically, when the sampling interval is sufficiently shorter than the Nyquist limit (i.e., less than half the period of the oscillation), these methods can yield reasonably accurate estimates. However, estimation bias increases substantially as the sampling interval approaches this limit. This

¹⁰Preliminary explorations in two extended contexts, reported in Sections S5 and S6 of the Supplementary Material, support the method's potential generality. In a predator–prey model (a different first-order system), the method achieved a 100% convergence rate and near-zero bias. In a bivariate coupled damped oscillator model (a more complex system involving 14 parameters), it showed a reduced convergence rate (80%) but successfully corrected bias for most parameters, highlighting both promise and areas for future refinement in highly complex models.

is demonstrated in our illustration with $\eta = -0.8$ and consistent with earlier simulation studies (e.g., Boker *et al.*, 2004). Therefore, researchers should be mindful of sampling density when applying these two-stage methods.

Two-stage estimation approaches, such as GLLA, provide practical approximations that can be informative in empirical studies. These methods have evolved over time, with developments such as obtaining derivative estimates using functional data analysis (Chow *et al.*, 2016; Trail *et al.*, 2014) or empirical Bayesian prediction (Deboeck, 2020). Another potential strategy to improve derivative approximation is to pre-process the data using interpolation techniques (e.g., B-spline imputation; Gasimova *et al.*, 2014), which may increase the effective sampling density and reduce bias in two-stage estimation.

However, it is important to recognize that parameter estimates from these two-stage methods may still be subject to bias. Unfortunately, this issue has received insufficient attention in the context of differential equation modeling. Several studies have noted and addressed the bias in the GLLA and LDE approaches. For instance, McKee *et al.* (2020) demonstrated that LDE produces biased estimates of the damping parameter in the presence of dynamic noise (process noise) and proposed a piecewise deterministic approximation algorithm to address this. Hunter (2016) highlighted that the derivative estimates from GLLA can be biased and exhibit correlated errors, and discussed the use of generalized orthogonal linear derivation (GOLD; Deboeck, 2010) to obtain derivatives with uncorrelated errors, though parameter estimates from GOLD may still remain biased.

The current study further addresses this gap by proposing an effective and adaptable bias-correction method for the GLLA approach, thereby contributing to the existing literature by offering a practical solution to improve the accuracy of parameter estimates. A key strength of this method is its robustness to various sources of approximation error in GLLA, such as those arising from sampling densities, the number of time points, and initial values. As demonstrated in our simulations, the SA algorithm effectively reduces estimation bias and yields accurate parameter estimates, regardless of the specific source of the initial approximation error. By highlighting the importance of bias correction in differential equation models, this study aims to encourage applied researchers to consider this issue in their analyses, promoting more rigorous practices. Furthermore, the ideas presented here may inspire future methodological research to explore further refinements of two-stage estimation approaches for differential equation models.

6.3. Limitations and future directions

Several limitations of the current study should be acknowledged. First, while the simulation study considered a range of conditions, it did not explore all possible scenarios that researchers might encounter in practice. For example, our investigation focused on relatively small damping parameter values ($\zeta = -0.04, 0, 0.04$) based on empirical findings in the psychological literature (Cho *et al.*, 2024). Examining a wider range of ζ values (e.g., Ollero *et al.*, 2025) would be a valuable direction for future research. Additionally, our simulation relied on a single observed indicator for the latent position, a setting in which separating measurement error from dynamic error can be particularly challenging. Extending the bias-correction method to models with multiple indicators represents an important avenue for further investigation.

Second, the proposed method relies on the injectivity of the mapping from the true parameter values to the expected value of the initial estimator (i.e., the GLLA estimator). The bias-correction procedure assumes local one-to-one behavior of this mapping so that the correction identifies a unique nearby parameter value. This condition is conceptually similar to the standard local identification requirement for parametric models (Hunter *et al.*, 2025), which typically involves the rank of the Jacobian of the model-implied moments. However, a key distinction is that our condition does not rely on the likelihood function, but rather on the expectation of the initial estimator. A sufficient condition for this local injectivity is that the Jacobian matrix of this mapping has full column rank in a neighborhood of the true parameter values. Although this assumption is plausibly satisfied in many typical applications, it

may not hold in all situations (e.g., when aliasing occurs due to oscillations beyond the Nyquist limit). In practice, this assumption can be empirically assessed by computing the numerical Jacobian at the obtained parameter estimates and checking its rank or condition number. Future work should formally establish sufficient conditions for local injectivity under realistic sampling constraints, to better identify the boundaries of the applicability of the bias-correction method.

Third, this study did not account for individual differences in model parameters. Beyond the average dynamics across persons, some researchers are also interested in how these dynamics vary across individuals, which can be modeled by incorporating random coefficients in differential equation models. In such models, the estimation of parameters becomes more complex due to the need to account for both fixed effects (representing the average dynamics) and random effects (capturing individual-level variations). As a result, correcting the bias of parameter estimation in these models can be challenging. Future research could explore how to apply the bias-correction method in differential equation models with random coefficients, investigating its utility in studying individual differences in dynamic processes.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/psy.2026.10104>.

Data availability statement. The data are available from the corresponding authors upon reasonable request.

Funding statement. The work of the first and second authors is supported by the National Natural Science Foundation of China (Grant No. 32471145, 32300938), the third author is supported by the National Natural Science Foundation of China (Grant No. 32571276), and the work of the second author is also supported by the Beijing Higher Education Innovation Center for Philosophy and Social Sciences.

Competing interests. The authors declare no competing interests.

References

- Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In S.-M. Chow, E. Ferrer, & F. Hsieh (Eds.), *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.
- Boker, S. M., Neale, M., & Rausch, J. (2004). Latent differential equation modeling with multivariate multi-occasion indicators. In K. V. Montfort, J. Oud, & A. Satorra (Eds.), *Recent developments on structural equation models: Theory and applications* (pp. 151–174). Springer.
- Boker, S. M., Tiberio, S. S., & Moulder, R. G. (2019). Robustness of time delay embedding to sampling interval misspecification. In K. V. Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 239–258). Springer.
- Chen, M., Chow, S.-M., & Hunter, M. D. (2018). Stochastic differential equation models with time-varying parameters. In K. V. Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 205–238). Springer.
- Cho, Y. W., Chow, S.-M., Marini, C. M., & Martire, L. M. (2024). Multilevel latent differential structural equation model with short time series and time-varying covariates: A comparison of frequentist and Bayesian estimators. *Multivariate Behavioral Research, 59*(5), 934–956.
- Chow, S.-M., Bendezú, J. J., Cole, P. M., & Ram, N. (2016). A comparison of two-stage approaches for fitting nonlinear ordinary differential equation models with mixed effects. *Multivariate Behavioral Research, 51*(2–3), 154–184.
- 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, 5*(2), 208–225.
- Deboeck, P. R. (2010). Estimating dynamical systems: Derivative estimation hints from sir Ronald A. Fisher. *Multivariate Behavioral Research, 45*(4), 725–745.
- Deboeck, P. R. (2020). Empirical Bayes derivative estimates. *Multivariate Behavioral Research, 55*(3), 382–404.
- Driver, C. C., Oud, J. H. L., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software, 77*, 1–35.
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*. (2nd ed.) Oxford University Press.
- Farr, R. H., Diamond, L. M., & Boker, S. M. (2014). Female same-sex sexuality from a dynamical systems perspective: Sexual desire, motivation, and behavior. *Archives of Sexual Behavior, 43*, 1477–1490.

- Gasimova, F., Robitzsch, A., Wilhelm, O., Boker, S. M., Hu, Y., & Hülür, G. (2014). Dynamical systems analysis applied to working memory data. *Frontiers in Psychology*, 5, 687.
- Gilbert, P., & Varadhan, R. (2016). *Numderiv: Accurate numerical derivatives*. [R package version 1.0-4].
- Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, 12(4), 748–762.
- Hu, Y., Boker, S., Neale, M., & Klump, K. L. (2014). Coupled latent differential equation with moderators: Simulation and application. *Psychological Methods*, 19(1), 56–71.
- Hu, Y., & Treinen, R. (2019). A one-step method for modelling longitudinal data with differential equations. *British Journal of Mathematical and Statistical Psychology*, 72(1), 38–60.
- Hunter, M. D. (2016). As good as GOLD: Gram-Schmidt orthogonalization by another name. *Psychometrika*, 81(4), 969–991.
- Hunter, M. D., Kirkpatrick, R. M., & Neale, M. C. (2025). Show me some ID: A universal identification program for structural equation models. *Psychometrika*, 90(2), 418–441.
- Leung, D. H. Y., & Wang, Y. G. (1998). Bias reduction using stochastic approximation. *Australian and New Zealand Journal of Statistics*, 40(1), 43–52.
- Lotka, A. J. (1920). Undamped oscillations derived from the law of mass action. *Journal of the American Chemical Society*, 42(8), 1595–1599.
- Luo, X., & Hu, Y. (2023). The competitive interaction between food cravings and unhealthy snacking: An application of the predator-prey model in psychology. *Current Psychology*, 42(32), 28484–28492.
- 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.
- Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81(2), 535–549.
- Ollero, M. J. F., Estrada, E., Hunter, M. D., & Cáncer, P. F. (2025). Characterizing affect dynamics with a damped linear oscillator model: Theoretical considerations and recommendations for individual-level applications. *Psychological Methods*, 30(5), 1095–1112.
- Ou, L., Hunter, M. D., & Chow, S.-M. (2019). What's for dynr: A package for linear and nonlinear dynamic modeling in R. *The R Journal*, 11(1), 91–111.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reed, R. G., Barnard, K., & Butler, E. A. (2015). Distinguishing emotional coregulation from codysregulation: An investigation of emotional dynamics and body weight in romantic couples. *Emotion*, 15(1), 45–60.
- Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics*, 22, 400–407.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedures. *The Annals of Mathematical Statistics*, 29(2), 373–405.
- Steele, J. S., & Ferrer, E. (2011). Latent differential equation modeling of self-regulatory and coregulatory affective processes. *Multivariate Behavioral Research*, 46(6), 956–984.
- Trail, J. B., Collins, L. M., Rivera, D. E., Li, R., Piper, M. E., & Baker, T. B. (2014). Functional data analysis for dynamical system identification of behavioral processes. *Psychological Methods*, 19(2), 175–187.
- Von Oertzen, T., & Boker, S. M. (2010). Time delay embedding increases estimation precision of models of intraindividual variability. *Psychometrika*, 75, 158–175.
- Zheng, Y., Xu, J., Li, K., & Hu, Y. (2024). A dynamical systems investigation of the co-regulation between perceived daily parental warmth and adolescent attention-deficit/hyperactivity disorder symptoms. *Research on Child and Adolescent Psychopathology*, 52(1), 111–124.

A. Appendix: Bias correction and SE estimation for multiple parameters

In this appendix, we establish key asymptotic properties of the bias-corrected estimator in multiparameter settings, extending the original result of Leung and Wang (1998) that was only applicable to single-parameter problems. Specifically, we demonstrate that the bias-corrected estimator achieves a bias of asymptotic order $O(n^{-1})$, aligning with that of the maximum likelihood estimator. Moreover, we also derive the asymptotic covariance matrix for the bias-corrected estimator and provide a practical procedure for standard error (SE) estimation.

A1. Theorem

Let p be the number of parameters, $(\theta^* \in \Theta \subseteq \mathbb{R}^p)$ denote the true parameter vector, and $(\hat{\theta} \in \Theta)$ denote the initial estimator that can be biased. Define $(L(\theta) = \mathbb{E}_{\theta} \hat{\theta})$ as the expected value of $(\hat{\theta})$ under the model parameterized by (θ) . The bias-correction equation is $L(\theta) - \hat{\theta} = 0$. The following regularity conditions are assumed for $\hat{\theta}$ and $L(\theta)$.

A1 (Positivity): For all $\theta \in \Theta$, the $p \times p$ Jacobian matrix $\dot{L}(\theta)$ exists. In addition, the diagonal elements of $\dot{L}(\theta^*)$, denoted by $\text{diag}(\dot{L}(\theta^*))$, are positive.

A2 (Lipschitz condition): There exist constants $K_1, K_2 > 0$ such that for all $\theta \in \Theta$,

$$K_1 \|\theta - \theta^*\|_2 \leq \|L(\theta) - L(\theta^*)\|_2 \leq K_2 \|\theta - \theta^*\|_2, \tag{A1}$$

where $\|x\|_2$ denotes the Euclidean norm for a vector x .

A3 (Taylor expansion): For all $\theta \in \Theta$, $L(\theta)$ admits the expansion:

$$L(\theta) = L(\theta^*) + \dot{L}(\theta^*)(\theta - \theta^*) + H(\theta, \theta^*) + \delta(\theta, \theta^*), \tag{A2}$$

in which $\dot{L}(\theta^*)$ is the Jacobian matrix, and $\delta(\theta) = o(\|\theta - \theta^*\|^2)$ uniformly in θ . The second-order term in the expansion is denoted as $H(\theta, \theta^*) = (H_1(\theta, \theta^*), \dots, H_p(\theta, \theta^*))^\top$, in which $H_j(\theta, \theta^*) = (\theta - \theta^*)^\top \ddot{L}_j(\theta^*)(\theta - \theta^*)/2$, $\ddot{L}_j(\theta^*)$ denotes the $p \times p$ Hessian matrix of $L_j(\theta^*)$, and $L_j(\theta^*)$ is the j th coordinate of $L(\theta^*)$ for each $j = 1, \dots, p$ (p is the number of parameters).

A4 (Stable initial estimator): Under the true model \mathbb{P}_{θ^*} , the covariance of initial estimator $\hat{\theta}$, denoted by $\Sigma_{\hat{\theta}}(\theta^*)$, is positive definite. Moreover, $\|\Sigma_{\hat{\theta}}(\theta^*)\|_2 = O(n^{-1})$, where $\|A\|_2$ stands for the spectral norm for a matrix A .

Under Assumptions A1–A4, the bias-corrected estimator θ_∞ satisfies

$$\|\mathbb{E}(\theta_\infty - \theta^*)\|_\infty = O(n^{-1}), \tag{A3}$$

in which $\|x\|_\infty$ denotes the maximum absolute value of the elements in x (i.e., the L^∞ norm of x).

A2. Proof of theorem

To establish Equation (A3), we define the conditional expectation error at the k th iteration as $e_k = \mathbb{E}(\theta_k - \theta^* | \hat{\theta})$, where $k = 1, \dots, K$, and K is the total number of iterations. The stochastic approximation (SA) algorithm iterates via $\theta_{k+1} = \theta_k - \gamma_k(\hat{\theta}_k - \hat{\theta})$. The subsequent conditional expectation error e_{k+1} is then given by

$$\begin{aligned} e_{k+1} &= \mathbb{E}(\theta_{k+1} - \theta^* | \hat{\theta}) = \mathbb{E}[\theta_k - \gamma_k(\hat{\theta}_k - \hat{\theta}) - \theta^* | \hat{\theta}] \\ &= \mathbb{E}[(\theta_k - \theta^*) - \gamma_k(\hat{\theta}_k - \hat{\theta}) | \hat{\theta}] = e_k - \gamma_k \mathbb{E}(\hat{\theta}_k - \hat{\theta} | \hat{\theta}) \\ &= e_k - \gamma_k \{ \mathbb{E}[L(\theta_k) | \hat{\theta}] - \hat{\theta} \}, \end{aligned}$$

in which the last equality follows from the conditional independence between $\hat{\theta}_k$ and $\hat{\theta}$ given θ_k . Substituting the Taylor expansion in Assumption A3, we derive the following recursive formula for the conditional expectation error:

$$\begin{aligned} e_{k+1} &= e_k - \gamma_k \dot{L}(\theta^*) \\ &= e_k - \frac{\gamma_k f_k}{2} - \gamma_k [L(\theta^*) - \hat{\theta} + \delta_k] \\ &= [I - \gamma_k \dot{L}(\theta^*)] e_k - \gamma_k \left(\frac{f_k}{2} + \delta_k \right) + \gamma_k [L(\theta^*) - \hat{\theta}], \end{aligned}$$

in which $f_k = \mathbb{E}[H(\theta_k, \theta^*) | \hat{\theta}]$ and $\delta_k = \mathbb{E}[\delta(\theta_k, \theta^*) | \hat{\theta}]$. $\gamma_k = \alpha k^{-\beta}$ is the step size at the k th iteration, where we choose $\alpha > 0$ and $\beta \in (1/2, 1]$. Taking expectation (with respect to $\hat{\theta}$) and noting that $\mathbb{E}\hat{\theta} = L(\theta^*)$, we obtain

$$\mathbb{E}e_{k+1} = [I - \gamma_k \dot{L}(\theta^*)] \mathbb{E}e_k - \gamma_k \mathbb{E}h_k, \tag{A4}$$

in which $h_k = f_k/2 + \delta_k$. Recursively applying Equation (A4) yields:

$$\mathbb{E}e_{k+1} = \prod_{i=1}^k [I - \gamma_i \dot{L}(\theta^*)] \mathbb{E}e_1 - \sum_{j=1}^k \left\{ \prod_{i=j+1}^k [I - \gamma_i \dot{L}(\theta^*)] \right\} \gamma_j \mathbb{E}h_j. \tag{A5}$$

The L^∞ norm of Equation (A5) can be bounded by the triangular equality:

$$\|\mathbb{E}e_{k+1}\|_\infty \leq \underbrace{\left\| \prod_{i=1}^k [I - \gamma_i \dot{L}(\theta^*)] \mathbb{E}e_1 \right\|_\infty}_{(I)} + \underbrace{\left\| \sum_{j=1}^k \left\{ \prod_{i=j+1}^k [I - \gamma_i \dot{L}(\theta^*)] \right\} \gamma_j \mathbb{E}h_j \right\|_\infty}_{(II)}. \tag{A6}$$

Let $\|A\|_\infty$ denote the maximum absolute value of the elements in a matrix A . To further bound (I) in Equation (A6), note that

$$\|I - \gamma_k \dot{L}(\theta^*)\|_\infty = \|\text{diag}(I - \gamma_k \dot{L}(\theta^*))\|_\infty = 1 - \alpha \gamma_k \tag{A7}$$

for sufficiently large k due to Assumption A1 and the fact that $\gamma_k \rightarrow 0$ as $k \rightarrow \infty$. The constant $\alpha > 0$ in Equation (A7) equals to the minimum diagonal element in $\dot{\mathbf{L}}(\boldsymbol{\theta}^*)$. Hence, there exists a positive integer k_0 such that for any $k \geq k_0$,

$$(I) \lesssim \|\mathbf{I} - \gamma_k \dot{\mathbf{L}}(\boldsymbol{\theta}^*)\|_\infty \|\mathbb{E} \mathbf{e}_1\|_1 \lesssim \prod_{i=k_0}^k (1 - \alpha \gamma_i) \|\mathbb{E} \mathbf{e}_1\|_1 \leq \exp\left(-\sum_{i=k_0}^k \alpha \gamma_i\right) \|\mathbb{E} \mathbf{e}_1\|_1, \tag{A8}$$

in which $\|\mathbf{x}\|_1$ denotes the L^1 norm of a vector \mathbf{x} , and $\mathbf{a} \lesssim \mathbf{b}$ means $\mathbf{a} \leq C\mathbf{b}$ with some $0 < C < \infty$. The last inequality in Equation (A8) follows from the fact that $1 - x \leq \exp(-x)$ for all $0 \leq x \leq 1$. Because $\sum_{i=k_0}^\infty \gamma_k = \infty$ (Equation (9)), the right-hand side of Equation (A8) is $o(1)$ as $k \rightarrow \infty$.

To bound (II) in Equation (A6), proceed similarly by noting that, for all $k \geq k_0$,

$$\begin{aligned} (II) &= \left\| \sum_{j=1}^k \left\{ \prod_{i=j+1}^k [\mathbf{I} - \gamma_i \dot{\mathbf{L}}(\boldsymbol{\theta}^*)] \right\} \gamma_j \mathbb{E} \mathbf{h}_j \right\|_\infty \\ &\leq \sum_{j=1}^k \left\| \left\{ \prod_{i=j+1}^k [\mathbf{I} - \gamma_i \dot{\mathbf{L}}(\boldsymbol{\theta}^*)] \right\} \gamma_j \mathbb{E} \mathbf{h}_j \right\|_\infty \\ &\leq \sum_{j=1}^k \left\| \left\{ \prod_{i=j+1}^k [\mathbf{I} - \gamma_i \dot{\mathbf{L}}(\boldsymbol{\theta}^*)] \right\} \right\|_\infty \gamma_j \|\mathbb{E} \mathbf{h}_j\|_1 \\ &\lesssim \sum_{j=k_0}^k \gamma_j \left\{ \prod_{i=j+1}^k (1 - \alpha \gamma_i) \right\} \|\mathbb{E} \mathbf{h}_j\|_1. \end{aligned} \tag{A9}$$

Now, we apply Lemma 1 of Sacks (1958) to conclude that

$$(II) \lesssim \|\mathbb{E} \mathbf{h}_\infty\|_1 \lesssim \|\boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*)\|_2, \tag{A10}$$

in which $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*) = \mathbb{E}[(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top]$.

To justify the last inequality in Equation (A10), recall that the L^1 norm of a vector is the sum of the absolute values of its components:

$$\|\mathbb{E} \mathbf{h}_\infty\|_1 = \sum_{j=1}^p |\mathbb{E}[(\mathbf{f}_\infty/2 + \boldsymbol{\delta}_\infty)_j]| \leq \frac{1}{2} \sum_{j=1}^p |\mathbb{E}[\mathbf{H}_j]| + \sum_{j=1}^p |\mathbb{E}[\boldsymbol{\delta}_j]|.$$

Thus, to bound $\|\mathbb{E} \mathbf{h}_\infty\|_1$, it suffices to bound each component $|\mathbb{E}[\mathbf{H}_j]|$ and $|\mathbb{E}[\boldsymbol{\delta}_j]|$ separately. We begin with the dominant term $|\mathbb{E}[\mathbf{H}_j]|$. By Assumption A3, $\mathbf{H}_j(\boldsymbol{\theta}_\infty, \boldsymbol{\theta}^*) = (\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top \dot{\mathbf{L}}_j(\boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)/2$. Then,

$$\begin{aligned} |\mathbb{E}[\mathbf{H}_j(\boldsymbol{\theta}_\infty, \boldsymbol{\theta}^*)]| &= \frac{1}{2} |\mathbb{E}[(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top \dot{\mathbf{L}}_j(\boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)]| \\ &= \frac{1}{2} |\mathbb{E}[\text{tr}((\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top \dot{\mathbf{L}}_j(\boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*))]| \\ &= \frac{1}{2} |\text{tr}(\dot{\mathbf{L}}_j(\boldsymbol{\theta}^*) \mathbb{E}[(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top])| \\ &= \frac{1}{2} |\text{tr}(\dot{\mathbf{L}}_j(\boldsymbol{\theta}^*) \boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*))| \\ &\leq \frac{1}{2} \|\dot{\mathbf{L}}_j(\boldsymbol{\theta}^*)\|_* \|\boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*)\|_2, \end{aligned}$$

where $\|\cdot\|_*$ denotes the nuclear (trace) norm and the inequality follows from Hölder's inequality for Schatten norms. Since $\dot{\mathbf{L}}_j(\boldsymbol{\theta}^*)$ is a fixed matrix (independent of n), its nuclear norm is bounded.

As for the remainder term $|\mathbb{E}[\boldsymbol{\delta}_j]|$, by Assumption A3, we have $\boldsymbol{\delta}_j(\boldsymbol{\theta}_\infty, \boldsymbol{\theta}^*) = o(\|\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*\|^2)$, which implies that $|\mathbb{E}[\boldsymbol{\delta}_j]|$ is of higher order and can be bounded by a constant multiple of $\|\boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*)\|_2$ using similar arguments. Combining the bounds for all components $j = 1, \dots, p$ yields the desired inequality.

To further analyze the order of $\|\boldsymbol{\Sigma}_{\boldsymbol{\theta}_\infty}(\boldsymbol{\theta}^*)\|_2$, consider the one-term expansion

$$\mathbf{L}(\boldsymbol{\theta}_\infty) = \mathbf{L}(\boldsymbol{\theta}^*) + \dot{\mathbf{L}}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*), \tag{A11}$$

in which $\hat{\boldsymbol{\theta}} = \kappa \boldsymbol{\theta}_\infty + (1 - \kappa) \boldsymbol{\theta}^*$ for some $0 \leq \kappa \leq 1$. Recall that $\mathbf{L}(\boldsymbol{\theta}_\infty) = \hat{\boldsymbol{\theta}}$ (see Equation (5)) and that $\mathbf{L}(\boldsymbol{\theta}^*) = \hat{\boldsymbol{\theta}}$. Therefore,

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}^*) = \mathbb{E}[(\hat{\boldsymbol{\theta}} - \mathbb{E} \hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \mathbb{E} \hat{\boldsymbol{\theta}})^\top] = \mathbb{E}[\dot{\mathbf{L}}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)(\boldsymbol{\theta}_\infty - \boldsymbol{\theta}^*)^\top \dot{\mathbf{L}}(\hat{\boldsymbol{\theta}})^\top]. \tag{A12}$$

By Assumptions A2 and A4, we have

$$\|\Sigma_{\theta_\infty}(\theta^*)\|_2 = \|\mathbb{E}[(\theta_\infty - \theta^*)(\theta_\infty - \theta^*)^\top]\|_2 \lesssim \|\Sigma_{\hat{\theta}}(\theta^*)\|_2 = O(n^{-1}). \quad (\text{A13})$$

Combining Equations (A6), (A8), (A10), and (A13) concludes the proof of the theorem.

A3. Standard error estimation

We next introduce a practical strategy to estimate the SEs of the bias-corrected estimator θ_∞ . More specifically, the theorem implies the consistency of θ_∞ ; therefore, by Equation (A11) and the continuous mapping theorem, we have

$$\theta_\infty - \theta^* = \dot{L}(\theta^*)^{-1}(\hat{\theta} - \mathbb{E}\hat{\theta}) + o_p(1). \quad (\text{A14})$$

Equation (A14) suggests that

$$\Sigma_{\theta_\infty}(\theta^*) \approx \dot{L}(\theta^*)^{-1} \Sigma_{\hat{\theta}}(\theta^*) \dot{L}(\theta^*)^{-\top}. \quad (\text{A15})$$

In practice, Equation (A15) cannot be directly evaluated due to the unknown θ^* . However, valid SEs can still be obtained using a consistent estimator of $\Sigma_{\theta_\infty}(\theta^*)$. For instance, we can consistently estimate $\dot{L}(\theta^*)$ by $\dot{L}(\theta_\infty)$ and $\Sigma_{\hat{\theta}}(\theta^*)$ by $\Sigma_{\hat{\theta}}(\theta_\infty)$. Monte Carlo (MC) estimates of \dot{L} and $\Sigma_{\hat{\theta}}$ can be substituted if closed-form expressions do not exist. For the simulation and empirical example reported in the current study, we approximate Equation (A15) by

$$\widehat{\Sigma}_{\theta_\infty}(\theta^*) = \left[\widehat{\dot{L}}(\bar{\theta}_K) \right]^{-1} \widehat{\Sigma}_{\hat{\theta}}(\bar{\theta}_K) \left[\widehat{\dot{L}}(\bar{\theta}_K) \right]^{-\top}, \quad (\text{A16})$$

in which $\bar{\theta}_K$ is computed by Equation (11), $\widehat{\Sigma}_{\hat{\theta}}(\bar{\theta}_K)$ is an MC estimate of $\Sigma_{\hat{\theta}}(\theta^*)$ using $\bar{\theta}_K$ as the data-generating parameters, and the Jacobian matrix $\widehat{\dot{L}}(\bar{\theta}_K)$ is estimated via numerical differentiation.