Daily Diary Data:

Effects of Cycles on Inferences

by

Yu Liu

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved January 2013 by the
Graduate Supervisory Committee:

Stephen G. West, Chair
Craig K. Enders
Mark Reiser

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

Daily dairies and other intensive measurement methods are increasingly used to study the relationships between two time varying variables X and Y. These data are commonly analyzed using longitudinal multilevel or bivariate growth curve models that allow for random effects of intercept (and sometimes also slope) but which do not address the effects of weekly cycles in the data. Three Monte Carlo studies investigated the impact of omitting the weekly cycles in daily dairy data under the multilevel model framework. In cases where cycles existed in both the time-varying predictor series (X) and the time-varying outcome series (Y) but were ignored, the effects of the within- and between-person components of X on Y tended to be biased, as were their corresponding standard errors. The direction and magnitude of the bias depended on the phase difference between the cycles in the two series. In cases where cycles existed in only one series but were ignored, the standard errors of the regression coefficients for the within- and between-person components of X tended to be biased, and the direction and magnitude of bias depended on which series contained cyclical components.

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my committee chair, Dr. Stephen G. West, for all his support towards my professional development. He mentored and nourished me with his precious experience and wisdom, both on this project and in my general academic pursuits. Without his guidance and persistent help this thesis would not have been possible.

I would like to thank my committee member Dr. Craig K. Enders for his feedback on the documents that helped me write more efficiently, and his personal communications that helped me become a more competent quantitative researcher.

I would like to thank my committee member Dr. Mark Reiser for providing relevant articles and useful suggestions.

I would like to thank Dr. Jenn-Yun Tein for graciously sharing her SAS syntax to generate multilevel data. She helped me understand macro writing in SAS, and made me a more competent programmer for the Monte Carlo simulations.

I would like to thank my dear friend Wan Yeung for supporting me through my difficult times, providing valuable insight into graduate school, for helping me organize the meetings.

I would like to thank my dear friend and colleague Ingrid Carlson Wurpts for her help with producing R graphics.

The most special thanks go to my parents and my fiancé for supporting me to pursue my professional goals.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

INTRODUCTION

In psychological research, participants are often asked to retrospect over a certain period of time (e.g. days, weeks, or months) to provide summary accounts of their psychological states and experiences (Bolger, Davis, & Rafaeli, 2003). However, most researchers agree that responses to retrospective questions tend to be biased by cognitive and motivational processes (e.g., McFarlane, Martin, & Williams, 1988). Among other important causes of bias, like recency and salience, is the participant's current state of mind (Reis & Gable, 2000). For example, summaries of mood over longer intervals tend to resemble the mood at the particular moment of measurement (Parkinson, Briner, Reynolds, & Totterdell, 1995, cited in Reis & Gable, 2000).

Diary data provide a good way to overcome such biases, characterizing ongoing experience with substantially greater accuracy than recollections (Reis & Gable, 2000). However, daily data gathered on individuals potentially often have a weekly, monthly, or seasonal cycle (Velicer & Molenaar, 2012). Despite the potential prevalence of cycles in phenomena, they are seldom directly modeled in the analyses. When the cyclical nature of the data is known, the cyclical effects can be statistically controlled or even removed. In economics, seasonal trends can be removed from the data prior to any time series analysis based on a priori information, or be controlled for by including as covariates variables sensitive to the same cyclical effect (Velicer & Molenaar, 2012). In analyses using techniques other than time series analysis, cycle-related artifacts and error variance can be controlled for by including parameters to estimate cyclical effects. For example, weekly cycles can be estimated using a dummy variable corresponding to weekday and

weekend (Gable & Reis, 1999), six dummy variables corresponding to days of the week (e.g., Armeli, Carney, Tennen, Affleck, & O'Neil, 2000), or Bowerman and O'Connell's (1979, cited in West & Hepworth, 1991) method of sine and cosine functions (e.g., Beal & Ghandour, 2011; Larsen & Kasimatis, 1990). However, when the cyclical nature of the data is not considered and thus not modeled, the presence of cycles in the data can potentially lead to problematic inferences. Little, if any, work has been done to systematically study the potential influence of ignoring cycles in the data on inferences from daily diary data.

Among the cycles that can potentially exist in the daily diary data, the weekly cycle is often of most interest because our work, schooling, and life usually follow a weekly schedule (Beal & Ghandour, 2011; Larsen & Kasimatis, 1990). Some empirical daily diary studies have taken this into account, and have modeled the weekly cycles (e.g., Armeli et al., 2000; Beal & Ghandour, 2011; Larsen & Kasimatis, 1990; Ram, Chow, Bowles, Wang, Grimm, Fujita, & Nesselroade, 2005); however, the majority of daily diary studies have ignored the possibility of cyclical patterns in the data.

The Profile of Weekly Cycles

Armeli et al. (2000, Figure 1, p. 983) presented profiles of average daily stress, alcohol consumption, and desire to drink by day of the week. These profiles indicated that in real data, weekly cycles can be monotonic, in which average daily levels only increase (or only decrease) by day of the week and then abruptly drop off (or go up) at some point of the weekly cycle, or they can be similar to the standard sine wave in which the average daily levels will increase for some days and decrease for some other days.

As mentioned above, besides using dummy variables, weekly cycles can be modeled using a combination of sine and cosine functions (e.g., Beal & Ghandour, 2011; Larsen & Kasimatis, 1990). This method is very flexible in that it can be used to form theoretical cycles, or to mimic empirical cycles having different periods (length of a complete cycle, as defined by Larsen, Augustine, & Prizmic, 2009), magnitudes (vertical distance between the peak and the nadir of one complete cycle), and specific shapes (profiles). To illustrate, take the Beal and Ghandour (2011) study as an example, which used the sine and cosine modeling approach to model the weekly cycle in the daily ratings of workplace affect collected from 65 employees for 21 days assessed when they left work. In their study, a 2-level multilevel model was fitted to the ratings of positive affect (PA) and the ratings of negative affect (NA), with Level-1 being measurement occasions, and Level-2 being persons. Their Level-1 model for positive affect with standardized regression coefficients was

$$PA_j = 0.01 \sin\left(\frac{2\pi * day}{7}\right) + 0.07 \cos\left(\frac{2\pi * day}{7}\right) + 0.10\ day + 0.01\ PA_{j-1} - 0.15\ Ike + e_j, \quad (1)$$

and their Level-1 model for negative affect with standardized regression coefficients was

$$NA_j = 0.00 \sin\left(\frac{2\pi * day}{7}\right) - 0.07 \cos\left(\frac{2\pi * day}{7}\right) - 0.01\ day + 0.11\ NA_{j-1} + 0.01\ Ike + e_j, \quad (2)$$

where j referred to today, $j - 1$ referred to yesterday; Ike was a dummy coded predictor representing the influence of Hurricane Ike[1], with values of 0 for all days prior to the landfall of Hurricane Ike and values of 1 for all days subsequent to it; day was the

[1] Hurricane Ike hit the area where the study was carried out during data collection, and it could be viewed as a negative affective event.

measure of time centered at the landfall of Hurricane Ike, not using whole number values for each day but rather the specific timestamps for each survey completion; $\sin(\frac{2\pi * day}{7})$ and $\cos\left(\frac{2\pi * day}{7}\right)$ referred to a sine function and a cosine function based on the day variable[2], respectively, and they together modeled the weekly cycle. The standardized regressions coefficients they reported were calculated using the following equation provided in Hox (2002), p.21,

$$standardized\ coefficient = \frac{(unstandardized\ coeff.)\times(stand.\ dev.\ explanatory\ var.)}{stand.\ dev.\ outcome\ variable}, \quad (3)$$

where the standard deviation of the outcome variable is defined as the standard deviation of the outcome variable when time is equal to zero, which is equivalent to the square root of the sum of the intercept variance and the residual variance obtained from the intercept-only model. For a time-varying predictor with only fixed effects (for example, the sine and cosine waves here), the standard deviation is calculated as the square root of the within-person variance.

Purposes of This Study

As described above, even though daily data gathered on individuals potentially often have cyclical patterns in the data, the majority of daily diary studies have ignored this possibility, and the potential influence of ignoring cycles on inference from repeated measures has not yet been systematically studied. The major goal of this study is to investigate (a) the effect of not modeling existing weekly cycles in daily diary data, and (b) factors that moderate this effect, under the longitudinal multilevel modeling

[2] For both PA and NA, the unstandardized regression coefficient of the sine function was non-significant, whereas the unstandardized regression coefficient of the cosine function was significant.

framework. I examined the impact of ignoring the cyclical patterns in three Monte Carlo studies.

To get a rough idea of the statistical models utilized in substantive research using the daily diary method, in December 2011, I examined the first 20 substantive research articles since Year 2000 returned by Google Scholar, using as the key words "daily diary method" and the exclusionary key word "event-based". Fifteen of these published studies utilized a longitudinal model (multilevel model or generalized estimation equation); of them, the majority included at least one continuous time-varying predictor at the within-person repeated measures level and at least one continuous time-invariant predictor at the between-person level.

Hence, in this study, I focused on the common two-level growth model: Level-1 models the within-person repeated measures, and Level-2 models the between-person differences. One time-varying predictor at the within-person repeated measures level and one time-invariant predictor at the between-person level were modeled to mimic the majority of published studies. I fixed the interval between two adjacent measurement occasions to 1 day, mimicking daily diary data, and fixed the lengths of the cycles to 1 week, mimicking weekly cycles. In the present study, I used the combination of sine and cosine waves to generate weekly cycles in which the average daily levels will increase for some days and decrease for some other days. I limited my focus to a time-balanced design, and assumed that the same cycle(s) characterize the responses of all the participants (i.e, there are only fixed effects of cycles, but no random effects).

Chapter 2

OVERVIEW OF METHODOLOGY, AND RESEARCH QUESTIONS

The Statistical Model for Data Generation

A longitudinal two-level model for the X series and a longitudinal two-level model for the Y series were used as the data-generating models for the simulation studies. I assumed that cycles existed in both of the series, and that the cycles could be characterized by the same combination of sine and cosine waves (i.e., the profile/shape of cycles was the same in both series). In addition, I assumed that neither series had a linear slope of time[3]. I first present the equations for the models, followed by a description of the data generation process.

The time-varying predictor. The uncentered Level-1 model for the time-varying predictor X can be written as

$$x_{it} = \pi_{0i} + \pi_{1i}\sin(\frac{2\pi t}{7}) + \pi_{2i}\cos(\frac{2\pi t}{7}) + e_{it}, \quad (4)$$

where x_{it} is the measure on the time-varying predictor of person i on day t, π_{0i} is the intercept for participant i, and π_{1i} and π_{2i} are the weights of the sine and cosine functions, respectively, on the time-varying predictor X for participant i. e_{it} is the Level-1 residual (within-person random error) for participant i on day t on the time-varying

[3] Some of the examined substantive research considered random linear slopes of time. I conducted a pilot simulation study where 100 people were measured daily for 63 consecutive days on a time-varying predictor and a outcome variable, both of which had a random linear slope of time. The results showed that not modeling the weekly cycles that existed in the two series did not have any impact on the average linear slopes or the slope residual variance. Indeed, the length of the weekly cycles (7 days) and the number of days that would provide sufficient power to detect the cycles (63 days, 9 weeks in the pilot study) made it unlikely for any effect of not modeling the cycles on the average linear slopes or the slope residual variance to occur. Thus, in the present studies, I omitted the linear slopes of time from the time-varying predictor series and the outcome series, and only assumed random intercepts for both series.

predictor X. That is to say, the within-person variability of the time-varying predictor X is captured by e_{it}, or, the within-person component of X is

$$xw_{it} = e_{it}. \tag{5}$$

Assuming a random effect for the intercept only, the Level-2 model for the time-varying predictor X can be written as

$$\begin{aligned} \pi_{0i} &= \mu_{00} + \zeta_{0i} \\ \pi_{1i} &= \mu_{10} \\ \pi_{2i} &= \mu_{20} \end{aligned}, \tag{6}$$

where μ_{00} is the population mean intercept of the time-varying predictor X, and μ_{10} and μ_{20} are the population mean regression weights of the sine and cosine functions. ζ_{0i} is the random effect (intercept residual) representing between-person variation in X. Put differently, the between-person component of X can be written as

$$xb_i = \zeta_{0i}. \tag{7}$$

The reduced mixed model form of Equations (4) and (6) can be written as

$$x_{it} = \mu_{00} + \mu_{10}\sin(\tfrac{2\pi t}{7}) + \mu_{20}\cos(\tfrac{2\pi t}{7}) + (\zeta_{0i} + e_{it}). \tag{8}$$

The outcome variable. I assumed that the within-person component of the time-varying predictor (xw_{it}) has an impact on the within-person part of the outcome Y, and that its between-person component (xb_i) has an impact on the between-person part of Y, where i represents person i, and t represents day t. How estimates of the within- and between-person components of the time-varying predictor X could be obtained is discussed in detail in the model fitting section below.

Suppose cycles in the time-varying predictor series and the outcome series are synchronized (when one takes on its highest value, the other is also at the peak). When

there is no autocorrelation in the residuals of the outcome series, the Level-1 model of the outcome measure can be written as

$$y_{it} = \beta_{0i} + \beta_{1i}\sin(\frac{2\pi t}{7}) + \beta_{2i}\cos(\frac{2\pi t}{7}) + \beta_{3i}xw_{it} + \varepsilon_{it}, \quad (9)$$

where y_{it} is the outcome measure of person i on day t, β_{0i} is the intercept for participant i, β_{1i} and β_{2i} are the weights of the sine and cosine functions for participant i, β_{3i} is the regression weight of the within-person part of the time-varying predictor X for participant i, and ε_{it} is the Level-1 residual (within-person random error) for participant i on day t, on the outcome variable y. Assuming a random effect for the intercept only, the Level-2 model can be written as:

$$\begin{array}{l} \beta_{0i} = \gamma_{00} + \gamma_{01}z_i + \gamma_{02}xb_i + \xi_{0i} \\ \beta_{1i} = \gamma_{10} \\ \beta_{2i} = \gamma_{20} \\ \beta_{3i} = \gamma_{30} \end{array}, \quad (10)$$

where Z is the between-person predictor (time-invariant), xb_i is the between-person component of the time-varying predictor X, γ_{00} to γ_{30} are the population mean regression weights, γ_{01} and γ_{02} represent change in the outcome variable Y for a 1-unit increase in the time-invariant predictor Z and for a 1-unit increase in the between-person component of the time-varying predictor X, respectively, and ξ_{0i} is the random effect (intercept residual) representing between-person variation in the outcome variable Y. The reduced mixed model form of Equations (9) and (10) can be written as

$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi t}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi t}{7}\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}). \quad (11)$$

When there is lag-1 autocorrelation [AR(1)] in the residuals of the outcome series, the model for the time-varying predictor stays the same, the Level-2 model of the

outcome variable stays the same, but the Level-1 residual term ε_{it} of the outcome variable now satisfies (see Congdon, 2006, p. 390):

$$\begin{aligned} &\varepsilon_{it} = \phi\varepsilon_{i(t-1)} + u_{it} \\ &\varepsilon_{i(t_0)} \sim N(0, \sigma^2) \\ &u_{it} \sim N(0, (1-\phi^2)\sigma^2) \end{aligned}, \tag{12}$$

where ϕ is the autoregressive correlation coefficient, and $\varepsilon_{i(t_0)}$ is the Level-1 residual of person i on the first day in the simulated data. Thus, for the first day, Equation (9) can still be used to describe the mixed model of the outcome variable:

$$y_{i(t_0)} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi(t_0)}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi(t_0)}{7}\right) + \gamma_{30}xw_{i(t_0)} + \gamma_{01}z_i + \gamma_{02}xb_i + \xi_{0i} + \varepsilon_{i(t_0)}, \tag{13}$$

whereas for all other days ($t > t_0$ in the simulated data), the mixed model of the outcome variable should be written as

$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi t}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi t}{7}\right) + \gamma_{30}zw_{it} + \gamma_{01}x_i + \gamma_{02}zb_i + \xi_{0i} + (\phi\varepsilon_{i(t-1)} + u_{it}). \tag{14}$$

Data Generation Process

I used SAS 9.2 to generate 1,000 datasets within each of the design cells. For the X series, I used the RANNOR function to generate a random normal variable that was then multiplied by the square root of the intercept residual variance in the X series, to form the random intercept residuals in the multilevel growth model of X (or equivalently, xb_i). For each time point, I used the RANNOR function to generate a random normal variable that was then multiplied by the square root of the Level-1 residual variance in X to form the Level-1 residual at that time point (or equivalently, xw_{it}), and used the built-in sin() and cos() functions to generate the sine and cosine waves. Then these values, as

well as the chosen values for the regression coefficients, were substituted into Equation (8) to generate the X series.

For the Y series, I first used the RANNOR function to generate the time-invariant predictor Z and a random normal variable that was then multiplied by the square root of the intercept residual variance in the Y series, to form the random intercept residuals in the multilevel growth model of Y. For each time point, I used the RANNOR function to generate a random normal variable that was then multiplied by the square root of the Level-1 residual variance in Y to form represents the Level-1 residual at that time point, and used the built-in sin() and cos() functions to generate the sine and cosine waves (taking into account the difference in phase between the cycles in X and the cycles in Y). When there was no autocorrelation in the residuals of the outcome series Y, these values and the chosen regression coefficient values were substituted into Equation (11) to generate the Y series. When there was lag-1 autocorrelation [AR(1)] in the residuals of the outcome series, the Level-1 residuals in Y were generated according to Equation (12), and then the generated predictor values, the chosen regression coefficient values, and the generated residual values were substituted into Equations (13) and (14) to generate the Y series. The computer script used to generate the data is presented in Appendix A.

Model Fitting Strategies

Disaggregating the Within- and Between-Person Components of the Time-Varying Predictor. The generated datasets were analyzed in SAS 9.2. Since it is assumed that the within-person component of the time-varying predictor (xw_{it}) has an impact on the within-person part of the outcome Y, and that its between-person component (xb_i) has an impact on the between-person part of Y, in this study the model

fitting has two steps. Step 1 involves disaggregating the time-varying predictor X into a within-person component (which acts as a Level-1 predictor) and a between-person component (which acts as an additional Level-2 predictor besides the time-invariant predictor Z), and Step 2 involves fitting a 2-level model for the outcome variable Y.

The standard method for disaggregating the within- and between-person effects of the time-varying predictor on the outcome, i.e. the person-mean centering strategy (Kreft, de Leeuw, & Aiken, 1995; Enders & Tofighi, 2007), implicitly makes the assumption that the value of the time-varying predictor is independent of time. Curran and Bauer (2011) realized that sometimes the time-varying predictor can have a linear slope of time. In such cases, they found that the person-mean centering strategy could not sufficiently disaggregate the within- and between-person effects of the time-varying predictor on the outcome, which in turn could lead to bias in parameter estimates. However, the possibility of the existence of cycles in both the time-varying predictor series and the outcome series was not considered in Curran and Bauer (2011). As is demonstrated below, when cycles exist in the time-varying predictor, the person-mean centering strategy cannot sufficiently disaggregate the within- and between-person components.

The person-specific expected value of x_{it} as defined in Equation (8) is

$$E_i(x_{it}) = \mu_{00} + \mu_{10}E_i\left(\sin\left(\frac{2\pi t}{7}\right)\right) + \mu_{20}E_i\left(\cos\left(\frac{2\pi t}{7}\right)\right) + \zeta_{0i}. \qquad (15)$$

Rearranging Equation (15) and inserting sample estimates, we obtain

$$\hat{\zeta}_{0i} = \bar{x}_i - \hat{\mu}_{00} - \hat{\mu}_{10}\left(\overline{\sin\left(\frac{2\pi t}{7}\right)}\right) - \hat{\mu}_{20}\left(\overline{\cos\left(\frac{2\pi t}{7}\right)}\right), \qquad (16)$$

where $\hat{\zeta}_{0i}$ is an estimate of the between-person component of X, $xb_i = \zeta_{0i}$. It can be proved that when the number of measurement occasions (i.e. the number of days) is a

multiple of the length of the weekly cycle (i.e., 7 days), $\overline{\sin\left(\frac{2\pi t}{7}\right)} = 0$, and $\overline{\cos\left(\frac{2\pi t}{7}\right)} = 0$. As is described in greater detail in the method section, this is satisfied in the present studies, so Equation (16) reduces to

$$\hat{\zeta}_{0i} = \bar{x}_i - \hat{\mu}_{00}. \tag{17}$$

Continuing to the expression of the within-person component, we can express the person- and time-specific residual defined in Equation (8) as

$$\hat{e}_{it} = \hat{x}_{it} - \hat{\mu}_{00} - \hat{\mu}_{10}\sin\left(\frac{2\pi t}{7}\right) - \hat{\mu}_{20}\cos\left(\frac{2\pi t}{7}\right) - \hat{\zeta}_{0i}. \tag{18}$$

Inserting the sample estimate of ζ_{0i} defined in Equation (17), Equation (18) can be re-written as

$$\hat{e}_{it} = (\hat{x}_{it} - \bar{x}_i) - \left(\hat{\mu}_{10}\sin\left(\frac{2\pi t}{7}\right) + \hat{\mu}_{20}\cos\left(\frac{2\pi t}{7}\right)\right). \tag{19}$$

Now the first term contains the person-mean centered time-varying predictor X, but it is not an appropriate representation of the within-person component of X, as it fails to take into account the cyclical trend reflected in the second term.

Though Curran and Bauer (2011) did not utilize it in their study[4], they suggested that a plausible method of obtaining appropriate sample estimates of the within- and between-person components of the time-varying predictor would be fitting a multilevel growth model to the time-varying predictor and asking for the empirical Bayes estimates of ζ_{0i} and e_{it}. I used this method in the pilot simulation study mentioned above, and it produced unbiased parameter estimates, and proper coverage rate of the 95% confidence intervals (which implied proper standard errors). In this study, I continued to use the

[4] For the purpose of disaggregating the within- and between-person effects of the time-varying predictor on the outcome, they calculated the deviation of the time-varying predictor from the individual-specific regression line reflecting the growth of the time-varying predictor over time.

empirical Bayes estimates as *appropriate* sample estimates of the within- and between-person components of the time-varying predictor X.

Model Fitting Strategy 1: Modeling the Cycles. In this study, for each generated dataset, I had two model fitting strategies, where Strategy 1 estimated the cyclical patterns (the correct model), and Strategy 2 ignored the cyclical patterns (the misspecified model). As is mentioned above, in this study, Step 1 of model fitting involves obtaining sample estimates of the within- and between-person components of the time-varying predictor X, and Step 2 involves fitting a 2-level model for the outcome variable Y.

For Strategy 1 (when the cycles were modeled), in Step 1 of model fitting, I obtained the empirical Bayes estimates of the Level-1 residuals and the Level-2 intercept residuals from a multilevel growth model fitted to the time-varying predictor X (using PROC MIXED with DDFM = KenwardRoger), and took them as the sample estimates of the within- and between-person components. In Step 2, I fitted a multilevel growth model to the outcome variable Y (using PROC MIXED with DDFM = KenwardRoger), with the Level-1 predictors being the sine wave and the cosine wave to capture the cyclical patterns, the estimated within-person component of the time-varying predictor ($\widehat{xw}_{it}$), and the Level-2 predictors of the random intercept being the time-invariant predictor Z and the estimated between-person component of the time-varying predictor ($\widehat{xb}_i$).

Model Fitting Strategy 2: Not Modeling the Cycles. For Strategy 2, I ignored the possibility of existence of cycles in the data, mimicking most empirical studies using daily diary data. If cycles did not exist in the data, person-mean centering, which is the standard practice in empirical research, would be appropriate for getting sample estimates

of the within- and between-person components of the time-varying predictor X (for an analytic proof, see Curran & Bauer, 2011). Hence, when I ignored the possibility of existence of cycles in the data, I used person-mean centering in Step 1 of model fitting. In Step 2, I fitted a multilevel growth model to the outcome variable Y (using PROC MIXED with DDFM = KenwardRoger), with the only Level-1 predictor being the person-mean centered time-varying predictor ($\hat{x}_{it} - \bar{x}_i$), and the Level-2 predictors of the random intercept being the time-invariant predictor Z and the person means of the time-varying predictor ($\bar{x}_i$).

I examined the impact of ignoring existing weekly cycles in daily diary data in three Monte Carlo studies. Studies 1 and 2 focused on situations in which cycles existed in both the outcome series Y and the time-varying predictor series X. In Study 1, I manipulated the profile of cycles[5], the synchronization of the cycles in the X series and the Y series, the serial dependency in the Y series, and the magnitude of standardized regression coefficient for $\widehat{xw}_{it}$. In Study 2, I examined the impact of synchronization of the cycles in the X series and the Y series in greater detail, and whether this impact would be moderated by the serial dependency in the Y series. Study 3 probed situations where cycles only exist in one of the two series. In Study 3, I examined the impact of ignoring weekly cycles that existed only in the time-varying predictor series X or only in the outcome series Y, and whether this impact would be moderated by the serial dependency in the Y series.

[5] The manipulation was simultaneously applied to both the X series and the Y series. Thus, the profile of cycles was the same for the two series.

Chapter 3

STUDY 1

In Study 1, I assumed that cycles existed in both the outcome series and the time-varying predictor series, and that the profile/shape of cycles was the same for the time-varying predictor series X and the outcome series Y.

Method for Study 1

This simulation study used a 3 (Profile of Cycles: cosine wave only, sine wave only with half the magnitude in the cosine wave only condition, or both sine and cosine waves) $\times 2$ (Synchronization of the cycles in the two series: synchronized or 180 °out of phase) $\times 2$ (Serial Dependency in the outcome series: zero autocorrelation, or non-zero lag-1 autocorrelation) $\times 3$ (Magnitude of Standardized Regression Coefficient for xw_{it} (the within- person component of the time-varying predictor): 0, 0.2, or 0.5) factorial design to generate the data.

A total of 1000 replications were generated for each condition using SAS 9.2. Each dataset was then analyzed in SAS 9.2 using two model fitting strategies: (1) a strategy in which cycles were modeled, and serial dependency was modeled as appropriate according to the population model, and (2) a strategy in which cycles were *not* modeled, whereas serial dependency was modeled as appropriate according to the population model. Comparing results from analyses with cycles modeled versus not modeled will help the understanding of the impact of failure to model cycles that exist on inferences from repeated measures.

The details of each design factor and other details of the model for data generation used in Study 1 are described below with a justification of the values selected for the study, followed by a summary of key parameter values in the study.

Details of design factors

Profile of Cycles. A review of the daily diary literature indicated very few studies have considered the possibility of weekly cycles. One exception was Beal and Ghandour (2011), who used the sine and cosine modeling approach to model the weekly cycle in the daily ratings of workplace affect. I used values from this study to provide plausible population values for some of the key parameters in my study. In Study 1, I included a cosine wave only condition (with a weight of 0 of the sine function and a standardized weight of .07 of the cosine function in both series, mimicking the magnitudes in the Beal & Ghandour, 2011, study), a sine wave only condition (with a standardized weight of .035 of the sine function and a weight of 0 of the cosine function in both series) to study the effect of the magnitude of the standardized weight[6], and a combined cosine and sine waves condition (with a standardized weight of .035 of the sine function and a standardized weight of .07 of the cosine function in both series) to mimic situations in which the cycle does not have a standard profile (see West & Hepworth, 1991, for an example). In terms of Cohen's (1988) norms, these represent *very small* effect sizes for the cyclical effects, but effect sizes that characterize those of the Beal and Ghandour (2011) study.

[6] A sine wave can be easily transformed into a cosine wave, so the difference between the sine wave only condition and the cosine wave only condition is the magnitude of the non-zero standardized regression weights.

Synchronization of the time-varying predictor series and the outcome series. In Study 1, weekly cycles existed in both the time-varying predictor series and the outcome series. In this case, the synchronization of these two series could be a potentially influential moderator of the impact of ignoring the cycles in the data on the inferences from the repeated measures. In the Armeli et al. (2000) study, the weekly cycles for the desire-to-drink series and the alcohol consumption series were approximately synchronized (when one took on the highest value, the other was also at the peak), whereas the weekly cycles for the stress series and the desire-to-drink series were approximately 180 °out of phase (when one took on its highest value, the other was at its nadir). Consequently, in Study 1 I decided to have a synchronized cycles condition, where the cycles in the two series were in phase, and a non-synchronized condition, where the cycles in the two series were out of phase by 180 °(π).

When the cycles in the outcome series and the time-varying predictor series were synchronized, the appropriate statistical model for data generation was described above in the statistical model section. When the cycles in the outcome series and the time-varying predictor series were not synchronized but rather out of phase by 180˚ (π), Equation (11) can still be used to describe the population model for the time-varying predictor variable, but the model for the outcome variable needs to be written as

$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi t}{7} - \pi\right) + \gamma_{20}\cos\left(\frac{2\pi t}{7} - \pi\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}), \quad (20)$$

which can be simplified to

$$y_{it} = \gamma_{00} - \gamma_{10}\sin\left(\frac{2\pi t}{7}\right) - \gamma_{20}\cos\left(\frac{2\pi t}{7}\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}). \quad (21)$$

Autocorrelation of Level-1 residuals in the outcome series. Serial dependency, or autocorrelation of Level-1 residuals, means that the Level-1 residual error terms for adjacent observations may be correlated. Trends, cycles, and serial dependency are three sources of non-independence of residuals in temporally ordered data (West & Hepworth, 1991). In daily diary data, even if trends of time do not exist and serial dependency (or autocorrelation of residuals) is modeled, unmodeled cycles may still contribute to non-independence of the Level-1 residuals, and may lead to bias and inflated standard errors in the estimation of the Level-1 residual variances and even the Level-1 residual correlations (autocorrelations). Multilevel models typically assume that the variances of the residuals are constant around the regression lines at both levels 1 and 2, and one important source of heteroskedasticity in the Level-2 residuals is that the model may *not* be correctly specified at level 1 (West, Ryu, Kwok, & Cham, 2011). Thus, if the failure to model the cycles in the data has any influence on the Level-1 model, it may also have an effect on the Level-2 intercept residual variance. Therefore, it is of interest to examine whether the failure to model the cycles in the data will have any effect on the modeling of autocorrelation of residuals and other parameters (e.g., the Level-1 residual variance, the intercept residual variance, etc.) in the outcome series. In Study 1, I decided to have a baseline no autocorrelation condition and a non-zero lag-1 autocorrelation [AR(1)] condition. When there was lag-1 autocorrelation [AR(1)] in the residuals of the outcome series, the autoregressive correlation coefficient ϕ was set to .80, which is within the range of values commonly used in past simulation studies (e.g., Kwok et al., 2007; Murphy, Beretvas, & Pituch, 2011; Murphy & Pituch, 2009).

Magnitude of standardized regression coefficient for the within-person component of the time-varying predictor. I chose three values for the standardized regression coefficient for xw_{it}, the within-person component of the time-varying predictor: 0, .2, and .5. As is explained in greater detail below, the within-person component of the time-varying predictor is captured by the Level-1 residual in the time-varying predictor series. Using a transformation of Equation (3) (see section "Summary of the key parameter values" below for details), the unstandardized regression coefficient of the within-person component of the time-varying predictor on the outcome can be calculated.

Details of the population model for generating datasets

Number of participants. The median number of participants was 118 in the examined substantive research articles described above (the first 20 substantive research articles since Year 2000 returned by Google Scholar, using as the key words "daily diary method" and the exclusionary key word "event-based"). Based on this, the number of participants in the simulation study was set to 100.

Number of measurement occasions. From the examination of the substantive research articles described above, I found that the median number of measurement occasions per person was 18. This number is quite small; however, very few of these examined studies addressed the possibility of cycles in the data. In those studies that addressed this possibility, the number of measurement occasions were generally larger, to ensure enough power to detect the weekly cycles. To cite three examples, in a study examining how the within-person associations among stress, alcohol use, and desire to drink varied as a function of several factors, Armeli et al. (2000) used daily diary

methods to examine 88 regular drinkers over 60 days, modeling the weekly cycles in the data using dummy variables corresponding to each day of the week. In a study of the weekly cycles in the daily moods, Larsen and Kasimatis (1990) examined the daily mood reports of 74 undergraduates for 84 consecutive days. Ram et al. (2005) examined the weekly cycles in daily emotion ratings from 179 college students for 52 consecutive days. Based on these, in this simulation study the number of measurement occasions per person was fixed at 63 to ensure enough power to detect the weekly cycles. With the interval size being 1 day, I had 9 weeks (day 1 to day 63) represented in the data. I centered the time variable at day 32 such that the mean of time was 0. In this way, the value of the time variable t at day 1 was -31, the value of the time variable t at day 32 was 0, and the value of the time variable t at day 63 was 31.

The time-invariant predictor. From my examination of the 20 substantive research articles described above, I found that 15 of them performed a multilevel model, and among these 15 articles, 2 used only binary time-invariant predictor(s), 9 used only continuous predictor(s), and 4 used both. It seems that the use of continuous time-invariant predictor(s) is more common than binary time-invariant predictor(s) in multilevel models, and thus I decided to model a continuous time-invariant predictor. For simplicity, the time-invariant predictor followed a standard normal distribution.

Standardized effect size of the time-invariant predictor and the between-person component of the time-varying predictor. Raudenbush and Liu (2001) provided the equation for calculating the standardized effect size of a Level-2 predictor (dummy coded treatment assignment in their example) on the random effect of polynomial trends (intercept, linear trend, quadratic trend, etc.) in the longitudinal

multilevel models framework. Thus, in calculating the standardized effect size of the time-invariant predictor and the between-person component of the time-varying predictor on the random intercept and the random slope, I can make use of the following equation

$$\delta_{pj} = \frac{\gamma_{pj}}{\sqrt{\tau_{pp}}}, \quad (22)$$

where δ_{pj} is the standardized effect size for the effect of the Level-2 predictor j ($j = 1$ for the time-invariant predictor, $j = 2$ for the between-person component of the time-varying predictor) on the growth parameters ($p = 0$ for the intercept, $p = 1$ for the slope), γ_{pj} is the population mean change in the corresponding growth parameter for a 1-unit change in the Level-2 predictor j, and τ_{pp} is the population variance of the random effect of the growth parameter. To simplify the study, the effects of the time-invariant predictor and the between-person component of the time-varying predictor on the random intercept had a standardized effect size of .50, a medium effect size according to Cohen (1988), as extended to clustered designs by Raudenbush and Liu (2000; 2001). Given the values of δ_{pj} (set to .50 in this study) and τ_{pp} (described in greater detail below), the corresponding value of γ_{pj} can easily be calculated.

Level-1 residual variance in the two series. In both the time-varying predictor series and the outcome series, the within-person residuals were assumed to be normally distributed with mean equal to zero and variance equal to one (i.e., $\sigma_x^2 = 1$ in the time-varying predictor series, and $\sigma^2 = 1$ in the outcome series), which is the general practice in simulation studies within the longitudinal multilevel framework (e.g., Curran & Bauer, 2010; Kwok et al., 2007; Murphy & Pituch, 2009).

The Level-2 intercept residual variance of the time-varying predictor series and the outcome series. In both series, the intercept residuals were assumed to be normally distributed with mean set to 0 and variance (τ_{x00} in the time-varying predictor series, and τ_{00} in the outcome series) set to .20, consistent with parameter values used in past simulation studies (e.g., Kwok et al., 2007; Raudenbush & Liu, 2000).

Magnitude of the mean intercept in the two series. For simplicity, the mean intercept in both series was fixed to zero in all conditions, which is a common strategy used in past simulation studies (e.g., Murphy et al., 2011; Murphy & Pituch, 2009).

Estimation method

Estimation of parameters in the multilevel modeling framework is typically done using maximum likelihood. In this context, statisticians distinguish between two kinds of maximum likelihood functions, full information maximum likelihood (FIML) and restricted maximum likelihood (REML). In FIML, both the regression coefficients and the variance components are included in the likelihood function; in contrast, in REML, only the variance components are included in the likelihood function, and the regression coefficients are estimated conditioning on the maximum likelihood estimates of the variance components (Raudenbush & Bryk, 2002). FIML is computationally easier, and an overall chi-square test based on the likelihood can be used to compare two models that differ in the fixed part, but estimates of variance components are generally negatively biased in small samples (Hox, 2002). REML partially addresses the issue, producing less biased estimates of variance components in small samples, but the REML log-likelihood can only be used to compare nested models for the covariance, not for the fixed parts. In this study, I fitted every model of the outcome variable Y twice, once using REML and

once using FIML, and took the REML estimates of variance components and the FIML estimates of the fixed parts (fixed effects and corresponding standard errors) for final analyses. FIML estimates of standard errors potentially can be biased, but in the pilot simulation study described above (which had similar population models and same number of participants and number of observations as this study), no such problem was found when the model fitting strategy was to estimate the existing cycles. Using the REML estimates of variance components and the FIML estimates of the fixed parts should be appropriate in this study, as it mimics substantive research in which researchers wish to be able to construct likelihood ratio tests of the fixed effects. For the time-varying predictor X, in modeling strategy 1, I fitted a multilevel model using FIML to obtain empirical Bayes estimates of the Level-1 and Level-2 residuals (within- and between-person components of X). As a check, I picked two conditions and compared the REML versus FIML empirical Bayes estimates of these within- and between-person components of X; the results showed no difference.

Summary of the key parameter values

A simple transformation of Equation (3) results in

$$unstandardized\ coefficient = \frac{(standardized\ coeff.)\times(stand.\ dev.\ outcome\ variable)}{stand.\ dev.\ explanatory\ var.},(23)$$

which can be used to calculate the unstandardized regression coefficient of the sine and cosine functions and the within-person component of the time-varying predictor. The standard deviation of the outcome variable when time is equal to zero is $SD_y = \sqrt{\sigma^2 + \tau_{00}} = \sqrt{1 + .2} = \sqrt{1.2} = 1.0954$, and the standard deviation of the time-varying predictor when time is equal to zero is $SD_x = \sqrt{\sigma_x^2 + \tau_{x00}} = \sqrt{1 + .2} = \sqrt{1.2} = 1.0954$.

The *sample* standard deviations of the sine and cosine functions are both equal to .7128, and the standard deviation of the within-person component of the time-varying predictor ($xw_{it} = e_{it}$) is 1. Thus, in the models of both series, for the sine and cosine functions, when the standardized regression coefficient equals 0, the unstandardized regression coefficient is also 0; when the standardized regression coefficient equals .035, the unstandardized regression coefficient is $\frac{(.035)(1.0954)}{.7128} = .0538$; and when the standardized regression coefficient equals .07, the unstandardized regression coefficient is $\frac{(.07)(1.0954)}{.7128} = .1076$. In the model of the outcome variable, the unstandardized regression coefficient of the within-person component of the time-varying predictor equals 0 when the standardized regression coefficient is 0, equals $\frac{(.20)(1.0954)}{1} = .2191$ when the standardized regression coefficient is .20, and equals $\frac{(.50)(1.0954)}{1} = .5477$ when the standardized regression coefficient is .50.

A simple transformation of Equation (22) results in

$$\gamma_{pj} = \delta_{pj} \times \sqrt{\tau_{pp}}, \tag{24}$$

so γ_{01} and γ_{02}, which represent change in the outcome variable Y for a 1-unit increase in the time-invariant predictor Z and for a 1-unit increase in the between-person component of X, are both equal to $(.50)\left(\sqrt{.20}\right) = .2236$.

Table 1 summarizes the key parameter values in Study 1.

Results for Study 1

The results are presented in the following order: 1) bias of the estimates of the fixed effects (the mean intercept, the effect of the within-person component of the time-varying predictor, and the effects of the time-invariant predictor and the between-person

component of the time-varying predictor) and their corresponding standard errors, the Level-1 residual variance, the Level-1 residual correlations, and the variance of the random intercepts, 2) empirical Type I error rate of the tests of the effect of the within-person component of the time-varying predictor, and 3) coverage of the confidence intervals (CI) of the fixed effects.

The impact of failure to model the cycles in the data, along with the four design factors on the simple bias[7] of the estimates of the parameters of interest was examined under the between-subjects ANOVA framework. All ANOVAs were conducted using SAS PROC GLM. Study 1 had 36 conditions ($3 \times 2 \times 2 \times 3$) $\times 1000$ replications $\times 2$ model fitting strategies = 72000 records to be submitted for analysis, resulting in very high power for detecting trivial effect sizes. Thus, rather than focusing on the significance tests, I used $\eta^2 > .01$ (a small effect according to Cohen, 1988) as a criterion for effect sizes worthy of consideration, which has been used in some previous simulation studies (e.g., Krull & MacKinnon, 1999; Murphy et al., 2011; Murphy & Pituch, 2009). Despite the factor of Model Fitting Strategy being within-subjects, I analyzed the data as a between-subjects design, so that the impacts of all factors on the parameters of interest were put in a single metric and thus were comparable (Gottschall, West, & Enders, 2012; Kwok et al., 2007). This choice does lower the statistical power of the test of the within subject terms, but given the focus on η^2 and the large number of replications, this issue should not be a concern.

To allow for comparison of bias across different parameters (fixed effects and variances), the standardized bias was also calculated within each design cell.

[7] The simple bias, or the raw bias, is defined as the difference between a parameter estimate and the true population value of that parameter.

Standardized bias (StdB) is defined as the deviation of the mean parameter estimate within each design cell from the true population value, divided by the standard deviation of the estimates within that design cell obtained from the correctly specified model (Collins, Schafer, & Kam, 2001; Enders & Tofighi, 2008). I used $\widehat{SE}_{\gamma_{00}}$ to represent the sample estimate of the standard error of $\hat{\gamma}_{00}$, and symbols for the sample estimates of the standard errors of other estimated fixed effects were defined in the same manner. Relative bias (RB) was calculated for the estimated standard errors of the estimated fixed effects, as the deviation of the sample estimate of the standard error from the population value of the standard error within each design cell, divided by the population value. I defined the population value of the standard error as the average of the 1000 sample estimates of the standard error within that design cell obtained from the correctly specified model. Rules of thumb guidelines for interpreting RB values are values less than 5% represent trivial bias, values between 5% and 10% represent moderate bias, and values greater than 10% represent substantial bias (Flora & Curran, 2004).

The type I error rate was examined for those conditions in which the true parameter value of γ_{30} was equal to zero. The full factorial design included a total of 24 different conditions (3 Profile of Cycle $\times 2$ Synchronization $\times 2$ Serial Dependency $\times 2$ Model Fitting Strategy) under which the empirical type I error rates could be evaluated. The impact of Model Fitting Strategy on the type I error inflation for detecting γ_{30} was evaluated by comparing the empirical Type I error rate to the expected confidence interval from the binomial distribution. Since the nominal Type I error rate is .05, the standard error of the binomial distribution is $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(.05)(.95)}{1000}} = .00689$. Thus,

empirical Type I error rates that fell outside of the range $[.036, .064]$ were considered problematic.

In addition to bias, the empirical coverage rate of the 95% CI's was used to evaluate the accuracy of the estimates of the fixed effects. The empirical coverage rate was calculated as the proportion of CI's that covered the population value of the corresponding fixed effect. If the model is correctly specified and the estimation method is working well, the coverage rate for the obtained 95% CI's of the fixed effects should be close to the nominal coverage rate, 95%. Following Collins et al. (2001), I considered the estimation of a parameter problematic if the coverage of its 95% CI fell lower than 90%.

No problems were encountered in fitting the models; the estimation procedures converged in all simulation data sets.

Bias of $\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$. When the simple bias of $\hat{\gamma}_{00}$, $\hat{\gamma}_{01}$ and $\hat{\gamma}_{02}$ were the dependent variables, no η^2 of the five design factors or their interaction effects in the ANOVA model was larger than .01. For the simple bias of the key parameter of interest, $\hat{\gamma}_{30}$, the Synchronization Factor had an impact ($\eta^2 = .026$). When the cycles in both series were synchronized, $\hat{\gamma}_{30}$ was positively biased ($\overline{StdB_{\hat{\gamma}_{30},synchronized\ cycles}} = .130$), whereas when the cycles were 180 °out of phase, $\hat{\gamma}_{30}$ was negatively biased ($\overline{StdB_{\hat{\gamma}_{30},not\ synchronized\ cycles}} = -.215$). This was modified by a Model Fitting Strategy by Synchronization two-way interaction ($\eta^2 = .026$). Figure 1[8] shows simple bias of $\hat{\gamma}_{30}$ by Model Fitting Strategy and Synchronization, averaged over the remaining design cells.

[8] I initially used side-by-side box plots for all effects to examine the distribution of outliers across conditions. Given that the distribution of outliers varied little across conditions, line graphs with 95% confidence intervals were displayed instead for better illustration of the mean level differences.

When the cycles in the two series were modeled, $\hat{\gamma}_{30}$ was always close to its true parameter value. When the cycles were not modeled, $\hat{\gamma}_{30}$ was positively biased when the cycles in the two series were synchronized, and negatively biased when they were 180 ° out of phase.

Bias of $\widehat{SE}_{\gamma_{00}}$, $\widehat{SE}_{\gamma_{30}}$, $\widehat{SE}_{\gamma_{01}}$, and $\widehat{SE}_{\gamma_{02}}$. When the simple bias of $\widehat{SE}_{\gamma_{00}}$, $\widehat{SE}_{\gamma_{30}}$, and $\widehat{SE}_{\gamma_{01}}$ were the dependent variables, no η^2 of the five design factors or their interaction effects in the ANOVA model was larger than .01. For the simple bias of $\widehat{SE}_{\gamma_{02}}$, the Model Fitting Strategy Factor had an impact ($\eta^2 = .121$). $\widehat{SE}_{\gamma_{02}}$ was negatively biased when the cycles in both series were not modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}},not\ modeling\ cycles}} = -.077$, representing a moderate bias) as compared to when the cycles were modeled (by definition, $\overline{RB_{\widehat{SE}_{\gamma_{02}},modeling\ cycles}} = 0$).

Simple bias of $\hat{\tau}_{00}$ and $\hat{\sigma}^2$. When the simple bias of $\hat{\tau}_{00}$ and $\hat{\sigma}^2$ were the dependent variables, no η^2 of the five design factors or their interaction effects in the ANOVA model was larger than .01.

Type I error rate for $\hat{\gamma}_{30}$. Figure 2 shows the trellis dot plot of the empirical Type I error rates of $\hat{\gamma}_{30}$. When the cycles were modeled, the empirical Type I error rates of $\hat{\gamma}_{30}$ were all within the acceptable range. In contrast, when the cycles were not modeled, the empirical Type I error rates were higher than .064 (the upper limit of the acceptable range) in 6 out of 12 conditions. Looking at only conditions in which the cycles were not modeled, it is obvious that the empirical Type I error rates increased as the Profile of Cycles changed from sine wave only to cosine wave only, then to both cosine and sine waves. The empirical Type I error rates were all within the reasonable range when the

cycles only had a sine wave, whereas with other profiles of cycles, the empirical Type I error rates tended to be too high. Since the only difference between the sine wave only condition and the cosine wave only condition was the magnitude of the cyclical components, it seems that the magnitude of the cyclical components is playing a role here. Moreover, when the cycles in the two series were synchronized, the difference in the empirical Type I error rates of $\hat{\gamma}_{30}$ between the both cosine and sine waves condition and the cosine wave only condition was larger when the Level-1 residuals in the outcome were independent (> .03), as compared to when the Level-1 residuals in the outcome were Lag-1 autocorrelated (< .005). When the cycles in the two series were *not* synchronized, this pattern was reversed: the difference in the empirical Type I error rates of detecting γ_{30} between the both cosine and sine waves condition and the cosine wave only condition was *smaller* when the Level-1 residuals in the outcome were independent (< .015), as compared to when the Level-1 residuals in the outcome were Lag-1 autocorrelated (> .025). These patterns are consistent with the analysis of the simple bias of $\hat{\gamma}_{30}$.

Coverage of the CI's of the fixed effects. The coverage rate of the 95% CI's was equal to or above 90% for $\hat{\gamma}_{00}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$ in all conditions. For $\hat{\gamma}_{30}$, the coverage rate of the 95% CI's fell below 90% in three conditions. In these three conditions, cycles in both series were 180 °out of phase and were *not* modeled, the Profile of Cycles had both cosine and sine waves, the Standardized Regression Coefficient for xw_{it} was .20 in one condition and .50 in the other two, and the Level-1 residuals were independent in two conditions and Lag-1 autocorrelated in the other one.

Impact of misspecifying the Level-1 covariance structure

Thus far the results of Study 1 seem to suggest that not modeling the existing cycles could lead to problematic inferences. However, remembering that cycles and serial dependency are two of the three sources of non-independence in temporally ordered data (West & Hepworth, 1991), I wished to investigate two questions: (a) Whether the non-independence arising from not modeling the cycles would be lessened by specifying a more general Level-1 covariance matrix in the outcome, and (b) if so, what factors would moderate this effect. In Study 1, there are two types of Level-1 convariance structures in the outcome, independent and AR(1), both of which are popular choices among substantive researchers analyzing daily diary data. The independence structure is nested under the AR(1) structure, providing a chance to examine these questions. Hence I can rephrase my questions thus: (a) when the Level-1 residuals in the outcome series are independent in the population model, can the non-independence arising from not modeling the cycles be absorbed by a misspecified AR(1) Level-1 covariance structure (serial dependency)? And (b) If so, what factors moderate this effect?

To answer these questions, in conditions where the Level-1 residuals in the outcome series were *independent* in the population model, I adopted a third model fitting strategy in which the cycles in the two series were not modeled, and an AR(1) Level-1 covariance structure in the outcome series was misspecified. Then I compared the two model fitting strategies that left out the cyclical components[9]. If specifying a more general Level-1 covariance structure would have a positive effect on the parameter estimates that suffered from ignoring the existing cycles, one would expect a difference in results from these two model fitting strategies. The results showed that specifying a

[9] One (model fitting strategy 2) specified the correct independent Level-1 covariance matrix, whereas the other, model fitting strategy 3, misspecified an AR(1) structure.

more general AR(1) Level-1 covariance structure did not provide a remedy. The parameter estimates of interest, particularly those that were biased when existing cycles were not modeled, $\hat{\gamma}_{30}$, $\hat{\gamma}_{02}$, their corresponding standard errors, the residual intercept variance $\hat{\tau}_{00}$ and the Level-1 residual variance $\hat{\sigma}^2$, did not differ for these two model fitting strategies, nor was the impact of misspecifying the Level-1 covariance structure moderated by any factor.

Summary of results in Study 1

In sum, the estimates of fixed effects, their corresponding standard errors, and the Level-1 and Level-2 residual covariance matrix elements were acceptable when cycles in both series were modeled. When the cycles were *not* modeled, $\hat{\gamma}_{30}$ tended to be biased, and the direction of bias depended on whether the cycles in the two series were synchronized or not. Moreover, $\widehat{SE}_{\gamma_{02}}$ was underestimated when the cycles were not modeled. When the Level-1 residuals in the outcome series were independent in the population model, misspecifying an AR(1) Level-1 covariance structure did not seem to provide a remedy to the problem arising from not modeling the cycles.

Chapter 4

STUDY 2

In Study 1 I examined conditions in which the cycles in the two series were synchronized versus 180 °out of phase (when one took on its highest value, the other was at its nadir). To examine this factor in greater detail, in Study 2, I conducted additional simulations. For simplicity, I assumed that cycles existed in both the outcome series and the time-varying predictor series, and that the profiles of cycles were the same for both the outcome series and the time-varying predictor series, as was the case in Study 1.

Method for Study 2

Study 2 used a 7 (Phase Difference of the cycles in the two series: 0, 1, 2, 3, 4, 5, or 6 days) ×2 (Serial Dependency in the outcome series: zero autocorrelation, or non-zero lag-1 autocorrelation) factorial design to generate the data. As was in Study 1, a total of 1000 replications were generated for each condition using SAS 9.2. Each dataset was then analyzed in SAS 9.2 with cycles modeled versus not modeled. To simplify the design, I fixed the standardized regression coefficient for xw_{it}, the within-person component of the time-varying predictor, to be 0.5. The profile/shape of cycles was also fixed such that the cycles always had both cosine and sine components. Other details of the data generation model were the same as those in Study 1.

The Phase Difference (PD) of the cycles in the two series was operationalized as the number of days that the weekly cycles in the outcome lagged behind the weekly cycles in the time-varying predictor. Hence Equation (8) can still be used to describe the time-varying predictor:

$$x_{it} = \mu_{00} + \mu_{10}\sin(\tfrac{2\pi t}{7}) + \mu_{20}\cos(\tfrac{2\pi t}{7}) + (\zeta_{0i} + e_{it}),$$

whereas the model for the outcome variable should be written as

$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi(t-PD)}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi(t-PD)}{7}\right) + \gamma_{30}xw_{it}$$
$$+\gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}). \quad (25)$$

When Phase Difference = 0, the cycles in the two series were synchronized. When Phase Difference = 1, the cycles in the outcome lagged 1 day behind those in the time-varying predictor, and so forth. Note that since the weekly cycles have a length of 7 days, Phase Difference = 6 (the weekly cycle in the outcome lagged 6 days behind the corresponding cycle in the time-varying predictor) can also be interpreted as the weekly cycle in the outcome is 1 day *ahead* of the *next* weekly cycle in the time-varying predictor. In this sense, Phase Difference only had 4 absolute values: 0, 1, 2, and 3 days. This factor was crossed with autocorrelation of Level-1 residuals in the outcome series (Serial Dependency), to see whether the impact of Phase Difference (if there is any) would interact with the existence of autocorrelation of Level-1 residuals in the outcome series.

Table 2 summarizes the key parameter values in Study 2.

Results for Study 2

No problems were encountered in fitting the models; the estimation procedures converged in all simulation data sets.

Bias of $\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$. When the simple bias of $\hat{\gamma}_{00}$ and $\hat{\gamma}_{01}$were the dependent variables, no η^2 of the three design factors or their interaction effects in the ANOVA model was larger than .01. Results for the simple bias of $\hat{\gamma}_{30}$ were more complex. Model Fitting Strategy had an impact ($\eta^2 = .018$), $\hat{\gamma}_{30}$ was close to its true

parameter value when the cycles in both series were modeled ($\overline{StdB_{\hat{\gamma}_{30},modeling\ cycles}} = .004$), but negatively biased when the cycles were not modeled ($\overline{StdB_{\hat{\gamma}_{30},not\ modeling\ cycles}} = -.295$). Phase Difference also had an impact ($\eta^2 = .032$), the smallest and greatest values (0, 1, and 6) of the phase difference between the weekly cycles in the two series led to less biased $\hat{\gamma}_{30}$ ($\overline{StdB_{\hat{\gamma}_{30}}}$ values were .102, .013, and .047, respectively), whereas phase difference values of 2, 3, 4, and 5 days (or the larger absolute values of Phase Difference) led to more biased $\hat{\gamma}_{30}$ ($\overline{StdB_{\hat{\gamma}_{30}}}$ values ranged from -.164 to -.423). These effects were modified by a Model Fitting Strategy by Phase Difference two-way interaction ($\eta^2 = .030$). To illustrate this pattern, Figure 3 displays simple bias of $\hat{\gamma}_{30}$ by Model Fitting Strategy and Phase Difference. The mean simple bias of $\hat{\gamma}_{30}$ was always close to zero when the cycles were modeled. When the cycles were not modeled, however, the mean simple bias of $\hat{\gamma}_{30}$ depends on the absolute values of Phase Difference: Phase Difference value of 0 day (synchronized cycles in the two series) led to positive bias in $\hat{\gamma}_{30}$, absolute Phase Difference value of 1 day (or, raw Phase Difference values of 1, and 7 days) led to mean simple bias values that were close to zero, whereas absolute Phase Difference values of 2 and 3 days (or, raw Phase Difference values of 2, 3, 4, and 5 days) led to negative mean simple bias values.

When the simple bias of $\hat{\gamma}_{02}$was the dependent variable, Model Fitting Strategy had an impact ($\eta^2 = .011$). $\hat{\gamma}_{02}$ was close to its true population value when the cycles in both series were modeled ($\overline{StdB_{\hat{\gamma}_{02},modeling\ cycles}} = .002$), but not when cycles were *not* modeled ($\overline{StdB_{\hat{\gamma}_{02},not\ modeling\ cycles}} = .204$).

Simple bias of $\widehat{SE}_{\gamma_{00}}$, $\widehat{SE}_{\gamma_{30}}$, $\widehat{SE}_{\gamma_{01}}$, and $\widehat{SE}_{\gamma_{02}}$. When the simple bias of $\widehat{SE}_{\gamma_{00}}$ and $\widehat{SE}_{\gamma_{01}}$ were the dependent variables, no η^2 of the three design factors or their interaction effects in the ANOVA model was larger than .01. Results for the simple bias of $\widehat{SE}_{\gamma_{30}}$ were more complex. Model Fitting Strategy had an impact ($\eta^2 = .012$): $\widehat{SE}_{\gamma_{30}}$ was inflated biased when the cycles in both series were not modeled ($\overline{RB_{\widehat{SE}_{\gamma_{30}},not\ modeling\ cycles}} = .004$) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{30}},modeling\ cycles}} = 0$ by definition). Phase Difference also had an impact ($\eta^2 = .016$): On average, the smallest and largest values of the phase difference between the weekly cycles in the two series led to $\widehat{SE}_{\gamma_{30}}$ values that were least biased ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}}$ all had an absolute value less than .001), whereas phase difference values of 2, 3, 4, and 5 days (or the larger absolute values of Phase Difference) led to inflated values of $\widehat{SE}_{\gamma_{30}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}}$ values were .003, .005, .005, and .003, respectively). Beyond these results, a Model Fitting Strategy by Phase Difference two-way interaction was found ($\eta^2 = .016$). Figure 4 shows simple bias of $\widehat{SE}_{\gamma_{30}}$ by Model Fitting Strategy and Phase Difference, averaged over the remaining design cells. When the cycles were modeled, the average relative bias of $\widehat{SE}_{\gamma_{30}}$ was 0 by definition. When the cycles were not modeled, $\widehat{SE}_{\gamma_{30}}$ tended to deviate from its true parameter value (defined as the average of the estimates of $SE_{\gamma_{30}}$ within a design cell obtained from the correctly specified models), and the amount of deviation depended on the phase difference between the weekly cycles in the two series: absolute Phase Difference values of 0 day (i.e., synchronized cycles in the two series) led to slightly underestimated $SE_{\gamma_{30}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = -.0014$), absolute Phase

Difference values of 1 day (or, raw Phase Difference values of 1, and 7 days) led to $\widehat{SE}_{\gamma_{30}}$ values that were rather close to the true parameter value ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = .0008$ for both conditions), whereas absolute Phase Difference values of 2 and 3 days (or, raw Phase Difference values of 2, 3, 4, and 5 days) led to over-estimation of $SE_{\gamma_{30}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}}$ values ranged from .0056 to .0097)[10].

For the simple bias of $\widehat{SE}_{\gamma_{02}}$, Model Fitting Strategy had an impact ($\eta^2 = .121$). $\widehat{SE}_{\gamma_{02}}$ was negatively biased when the cycles in both series were not modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}},not\ modeling\ cycles}} = -.077$, representing a moderate bias) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}},modeling\ cycles}} = 0$ by definition).

Simple bias of $\hat{\tau}_{00}$ and $\hat{\sigma}^2$. When the simple bias of $\hat{\tau}_{00}$ was the dependent variable, the η^2 of the three design factors and their interaction effects in the ANOVA model were all smaller than .01. When the simple bias of $\hat{\sigma}^2$ was the dependent variable, Model Fitting Strategy had an impact ($\eta^2 = .021$). On average, not modeling the cycles in the two series led to relatively larger overestimation of σ^2 ($\overline{StdB_{\hat{\sigma}^2,not\ modeling\ cycles}} = .396$), as compared to modeling the cycles ($\overline{StdB_{\hat{\sigma}^2,modeling\ cycles}} = .016$). Phase Difference also had an impact ($\eta^2 = .010$): On average, the smallest and largest values of the phase difference between the weekly cycles in the two series led to a $\hat{\sigma}^2$ that was close to the population value ($\overline{StdB_{\hat{\sigma}^2}}$ values ranged from .006 to .115), whereas phase difference values of 2, 3, 4, and 5 days led to relatively larger overestimation of σ^2 ($\overline{StdB_{\hat{\sigma}^2}}$ values ranged from .241 to .386).

[10] These mean RB values all represented trivial bias in $\widehat{SE}_{\gamma_{30}}$ according to Flora and Curran (2004).

Coverage of the CI's of the fixed effects. The coverage rate of the 95% CI's was equal to or above 90% for $\hat{\gamma}_{00}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$ in all conditions. For $\hat{\gamma}_{30}$, the coverage rate of the 95% CI's fell below 90% in two conditions. In these two conditions, cycles in both series were not modeled, the Level-1 residuals in the outcome series were independent, and Phase Difference took on the largest absolute value in the design (i.e., 3 days, which corresponds to raw Phase Difference values of 3 and 4 days).

Impact of misspecifying the Level-1 covariance structure

In Study 2, to investigate whether the non-independence arising from not modeling the cycles would be lessened by specifying a more general Level-1 covariance matrix in the outcome, I again adopted the third model fitting strategy[11] in conditions where the Level-1 residuals in the outcome series were *independent* in the population model. Then I compared the two model fitting strategies that left out the cyclical components[12], since if specifying a more general Level-1 covariance structure could improve parameter estimates that suffered from not modeling the cycles, there should be some difference in results from these two model fitting strategies. The results showed that specifying a more general AR(1) Level-1 covariance structure did not improve the problematic parameter estimates obtained from model fitting strategy 2. The parameter estimates of interest, particularly those that were biased when existing cycles were not modeled, $\hat{\gamma}_{30}$, $\hat{\gamma}_{02}$, the corresponding standard errors, the residual intercept variance $\hat{\tau}_{00}$ and the Level-1 residual variance $\hat{\sigma}^2$, did not differ for these two model fitting strategies,

[11] That is, the cycles in the two series were not modeled, and an AR(1) Level-1 covariance structure in the outcome series was misspecified.

[12] One (model fitting strategy 2) specified the correct independent Level-1 covariance matrix, whereas the other, model fitting strategy 3, misspecified an AR(1) structure.

nor was the impact of misspecifying the Level-1 covariance structure moderated by any factor.

Summary of results in Study 2

As in Study 1, the estimates of fixed effects, their corresponding standard errors, and the Level-1 and Level-2 residual covariance matrix elements were acceptable when cycles in both series were modeled. When the cycles were *not* modeled, $\hat{\gamma}_{30}$ tended to be biased, and the direction and magnitude of bias depended on the absolute value of the Phase Difference between the cycles in the two series; $\hat{\gamma}_{02}$ was positively biased, $SE_{\gamma_{30}}$tended to be underestimated when the cycles in the two series were synchronized, but overestimated when the cycles were not synchronized, and the magnitude of overestimation depended on the absolute value of the Phase Difference; $SE_{\gamma_{02}}$ tended to be underestimated, and σ^2 tended to be overestimated. When the Level-1 residuals in the outcome series were independent in the population model, specifying an AR(1) Level-1 covariance structure did not seem to provide a remedy to the problem arising from not modeling the cycles.

Chapter 5

STUDY 3

In Studies 1 and 2, cycles existed in both of the series. Realizing that there could be situations in real life where cycles only exist in one of the two series, I conducted additional simulations in Study 3 to investigate the impact of not modeling cycles that exist in only one of the two series.

Method for Study 3

Study 3 used a 2 (Existence of Cycles: only in outcome, or only in time-varying predictor) × 2 (Serial Dependency in the outcome series: zero autocorrelation, or non-zero lag-1 autocorrelation) factorial design to generate the data. A total of 1000 replications were generated for each condition using SAS 9.2. Each dataset was then analyzed in SAS 9.2 with cycles modeled versus not modeled. To simplify the design, I fixed the standardized regression coefficient for xw_{it}, the within-person component of the time-varying predictor, to be 0.5. The profile/shape of cycles was also fixed such that the cycles always had both cosine and sine components. Other details of the data generation model were the same as those in Study 1.

When cycles only existed in the time-varying predictor, model fitting strategies 1 and 2 stayed the same (i.e., empirical Bayes estimates were used to represent the within-person component and the between-person component of time-varying predictor when cycles were modeled, whereas the person-mean centering strategy was used when cycles were *not* modeled). When cycles only existed in the outcome series, however, for both model fitting strategies, person-mean centering was used to disaggregate the time-varying predictor. This is because in this scenario, getting empirical Bayes estimates and person-

mean centering are essentially equivalent ways of obtaining sample estimates of the within- and between person components of the time-varying predictor. Thus, when cycles only existed in the outcome series, the only difference between model fitting strategies 1 and 2 would be whether cycles were modeled in the outcome series.

Table 3 summarizes the key parameter values in Study 3.

Results for Study 3

No problems were encountered in fitting the models; the estimation procedures converged in all simulation data sets.

Simple bias of $\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$. When the simple bias of $\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$ and $\hat{\gamma}_{02}$ were the dependent variables, no η^2 of the three design factors or their interaction effects in the ANOVA model was larger than .01.

Simple bias of $\widehat{SE}_{\gamma_{00}}$, $\widehat{SE}_{\gamma_{30}}$, $\widehat{SE}_{\gamma_{01}}$, and $\widehat{SE}_{\gamma_{02}}$. When the simple bias of $\widehat{SE}_{\gamma_{00}}$ and $\widehat{SE}_{\gamma_{01}}$ was the dependent variable, no η^2 of the three design factors or their interaction effects in the ANOVA model was larger than .01. For the simple bias of $\widehat{SE}_{\gamma_{30}}$, Existence of Cycles had an impact ($\eta^2 = .013$): On average $SE_{\gamma_{30}}$ was overestimated when cycles only existed in the outcome series Y ($\overline{RB_{\widehat{SE}_{\gamma_{30}},cycles\ only\ in\ Y}} = .003$), but was underestimated when cycles only existed in the time-varying predictor series X ($\overline{RB_{\widehat{SE}_{\gamma_{30}},cycles\ only\ in\ X}} = -.001$). However, this effect was moderated by the Model Fitting Strategy Factor ($\eta^2 = .013$), which is portrayed in Figure 5. When cycles only existed in the outcome series, ignoring the cycles led to overestimated $SE_{\gamma_{30}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = .005$, representing trivial bias) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = 0$ by definition). When cycles only existed in the time-varying predictor,

ignoring the cycles led to slightly underestimated $SE_{\gamma_{30}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = -.001$, representing trivial bias) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{30}}}} = 0$ by definition).

For the simple bias of $\widehat{SE}_{\gamma_{02}}$, Model Fitting Strategy had an impact ($\eta^2 = .033$): On average $SE_{\gamma_{02}}$ was underestimated when the cycles were not modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}},not\ modeling\ cycles}} = -.038$) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}},modeling\ cycles}} = 0$ by definition). Existence of Cycles also had an impact ($\eta^2 = .033$): On average $\widehat{SE}_{\gamma_{02}}$ was close to the true parameter value when cycles only existed in the outcome series Y ($\overline{RB_{\widehat{SE}_{\gamma_{02}},cycles\ only\ in\ Y}} = .000$), but underestimated when cycles only existed in the time-varying predictor series X ($\overline{RB_{\widehat{SE}_{\gamma_{02}},cycles\ only\ in\ X}} = -.038$). These effects were modified by a Model Fitting Strategy by Existence of Cycles interaction ($\eta^2 = .033$), which is portrayed in Figure 6. When cycles only existed in Y, on average values of $\widehat{SE}_{\gamma_{02}}$ were identical whether the cycles were modeled or not. When cycles only existed in X, not modeling the cycles led to an underestimated $SE_{\gamma_{02}}$ ($\overline{RB_{\widehat{SE}_{\gamma_{02}}}} = -.076$, representing a moderate bias) as compared to when the cycles were modeled ($\overline{RB_{\widehat{SE}_{\gamma_{02}}}} = 0$ by definition).

Simple bias of $\hat{\tau}_{00}$ and $\hat{\sigma}^2$. When the simple bias of $\hat{\tau}_{00}$ and $\hat{\sigma}^2$ were the dependent variables, no η^2 of the three design factors or their interaction effects in the ANOVA model was larger than .01.

Coverage of the CI's of the fixed effects. The coverage rate of the 95% CI's was equal to or above 90% in all conditions for all the fixed effects of interest ($\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$).

Impact of misspecifying the Level-1 covariance structure

In Study 3, to investigate whether the non-independence arising from not modeling the cycles that existed in only one series would be lessened by specifying a more general Level-1 covariance matrix in the outcome, I adopted the third model fitting strategy[13] in conditions where the Level-1 residuals in the outcome series were *independent* in the population model. Then I compared the two model fitting strategies that left out the cyclical components[14], since if specifying a more general Level-1 covariance structure could provide even a partial remedy, results from these two model fitting strategies should be different. The results showed that specifying a more general AR(1) Level-1 covariance structure did not improve the problematic parameter estimates obtained from model fitting strategy 2. The parameter estimates of interest, particularly those that were biased when existing cycles were not modeled, $\hat{\gamma}_{30}$, $\hat{\gamma}_{02}$, the corresponding standard errors, the residual intercept variance $\hat{\tau}_{00}$ and the Level-1 residual variance $\hat{\sigma}^2$, did not differ for these two model fitting strategies, nor was the impact of misspecifying the Level-1 covariance structure moderated by any factor.

Summary of results in Study 3

In sum, the estimates of fixed effects, their corresponding standard errors, and the Level-1 and Level-2 residual covariance matrix elements were acceptable when cycles were modeled. When the cycles were *not* modeled, $\widehat{SE}_{\gamma_{30}}$ tended to be biased, and the direction of bias depended on which series had a cyclical component; $\widehat{SE}_{\gamma_{02}}$ was underestimated when cycles only existed in the time-varying predictor series, but when

[13] That is, the cycles in the two series were not modeled, and an AR(1) Level-1 covariance structure in the outcome series was misspecified.

[14] One (model fitting strategy 2) specified the correct independent Level-1 covariance matrix, whereas the other, model fitting strategy 3, misspecified an AR(1) structure.

cycles only existed in the outcome series, not modeling cycles did not lead to lower $\widehat{SE}_{\gamma_{02}}$. When the Level-1 residuals in the outcome series were independent in the population model, specifying an AR(1) Level-1 covariance structure did not seem to provide a partial remedy to the problem arising from not modeling the cycles.

Sandwich Estimator: Does It Take Care of Everything?

Some applied researchers believe that the sandwich estimator can substantially improve standard errors and possibly parameter estimates. To probe whether this belief is correct, I selected 3 conditions from the 3 Monte Carlo studies to examine whether the sandwich estimator can correct for the bias in parameter estimates or standard errors due to ignoring the existing weekly cycles.

From Study 1, I picked a condition where the cycles had sine and cosine waves, cycles in the X and Y series were out of phase by 180 °, the Level-1 residuals in Y were independent, and the standardized regression coefficient for $\widehat{xw}_{it}$ was .50 (Condition 1), and fitted models using the sandwich estimator when cycles were modeled *and* when cycles were not modeled. This condition was picked because when using the model-based estimator with DDFM = KenwardRoger, when cycles were modeled, parameter estimates and standard errors were proper; whereas when cycles were not modeled, this was one of the worst conditions for $\hat{\gamma}_{30}$ and for $\hat{\gamma}_{02}$: The coverage rate of the 95% CI's for $\hat{\gamma}_{30}$ was the lowest among the conditions in Study 1, and $\widehat{SE}_{\gamma_{02}}$ was underestimated.

From Study 2, I picked a condition in which the cycles in Y lagged 1 day behind the cycles in X, and the Level-1 residuals in Y were independent (Condition 2) , and fitted models using the sandwich estimator when cycles were *not* modeled. This condition was picked because when using the model-based estimator with DDFM = KenwardRoger,

when the cycles were not modeled, the parameter estimates were acceptable for $\hat{\gamma}_{30}$ (least biased fixed effect with smallest RB of the corresponding SE, as compared to other conditions in Study 2 where cycles were not modeled), but worst for $\hat{\gamma}_{02}$ in Study 2 (most biased fixed effect leading to lowest coverage rate of the 95% CI's, which was still above 90%).

From Study 3, I picked a condition where cycles only existed in the outcome series and the Level-1 residuals in Y were Lag-1 autocorrelated (Condition 3), and fitted models using the sandwich estimator when cycles were *not* modeled. This condition was picked because when using the model-based estimator with DDFM = KenwardRoger, when the cycles were not modeled, the parameter estimates were acceptable for $\hat{\gamma}_{02}$ (least biased fixed effect as compared to other conditions in Study 3 where cycles were not modeled, and $\widehat{SE}_{\gamma_{02}}$ was unbiased). $\hat{\gamma}_{30}$ was least biased as compared to other conditions in Study 3 where cycles were not modeled, but its corresponding SE had the greatest RB.

In running models using the sandwich estimator, I used Model Fitting Strategies 1 (modeling the cycles) and 2 (not modeling the cycles) for Condition 1, and only Model Fitting Strategy 2 for Conditions 2 and 3. I specified the EMPIRICAL option in PROC MIXED (to invoke the sandwich estimator) and chose DDFM = BETWITHIN. When the EMPIRICAL option is specified, the Satterthwaite and Kenward-Roger degrees of freedom methods are not available, and I had to choose among CONTAIN, BETWITHIN, and RESIDUAL. To mimic what most substantive researchers would do in their data analyses, I chose DDFM = BETWITHIN, which yielded the same degrees of freedom (97) as the widely used statistical program HLM6 for $\hat{\gamma}_{00}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$. Although the degrees of freedom that DDFM = BETWITHIN produced for $\hat{\gamma}_{30}$, 6199,

was *not* identical to that produced by HLM6, 6296, the difference was trivial given the very large degrees of freedom. Therefore, all other things being equal, these two degrees of freedom (6199 and 6296) should produce essentially the same *p* values. In contrast, DDFM = CONTAIN yielded df = 6198 for $\hat{\gamma}_{01}$ and $\hat{\gamma}_{02}$, which were much larger than those outputted by HLM6 (even though the df for $\hat{\gamma}_{00}$ (98) and for $\hat{\gamma}_{30}$ (6198) were acceptable). DDFM = RESIDUAL on the other hand yielded the same df (6296) for $\hat{\gamma}_{00}$, $\hat{\gamma}_{30}$, $\hat{\gamma}_{01}$, and $\hat{\gamma}_{02}$, which was acceptable for $\hat{\gamma}_{30}$, but much larger than those calculated by HLM6 for the other three parameters.

In all of the 3 conditions, on average the sandwich estimator led to smaller $\widehat{SE}_{\gamma_{30}}$ and $\widehat{SE}_{\gamma_{02}}$ as compared to their counterparts using the model-based estimator with DDFM = KenwardRoger, but did not produce different fixed effect estimates. As a result, with biased fixed effect estimates (though the bias was small in some conditions) and smaller corresponding standard errors, the coverage rates of the 95% CI's of $\hat{\gamma}_{30}$ and $\hat{\gamma}_{02}$ decreased slightly. For these 3 conditions, Table 4 summarizes the key parameters ($\hat{\gamma}_{30}$ and $\hat{\gamma}_{02}$) estimated using the empirical estimator, and compares them to the results using the model-based estimator with DDFM = KenwardRoger. For Conditions 2 and 3, results obtained from model fitting strategy 1 (modeling the cycles) using the model-based estimator with DDFM = KenwardRoger was also listed in the table to offer a comparison. Using the sandwich estimator does not seem to correct for biases arising from not modeling the existing cycles. Given this, producing smaller standard error estimates by using the sandwich estimator may not be very beneficial for $\hat{\gamma}_{30}$, since the failure to model cycles only led to trivial bias in $\widehat{SE}_{\gamma_{30}}$, and the use of the sandwich estimator did

not produce a substantial difference in $\widehat{SE}_{\gamma_{30}}$. For $\hat{\gamma}_{02}$, the use of the sandwich estimator when cycles were not modeled could even be harmful, because in Conditions 1 and 2 the failure to model cycles resulted in moderately underestimated $\widehat{SE}_{\gamma_{02}}$, and the use of the sandwich estimator in these two conditions added to the negative bias in $\widehat{SE}_{\gamma_{02}}$.

Chapter 6

DISCUSSION

The main purpose of this study was to investigate the effect of not modeling existing weekly cycles in daily diary data on the relationships between two time varying variables X (the time-varying predictor) and Y (the outcome), under the longitudinal multilevel modeling framework. Based on my population model, I expected that the within-person component of X would have an impact on the within-person part of the outcome Y, and that its between-person component would have an impact on the between-person part of Y. The target parameters included fixed effects, their corresponding standard errors, the within-subject residual variance, and the random intercept residual variance in the model of Y. The simulation results showed a general pattern of effects of ignoring existing cycles on the estimation of the fixed effects, their corresponding standard errors, and the within-subject residual variance. In cases where cycles existed in both the X series and the Y series, ignoring the cycles generally led to bias in the estimated effects of both the within- and between-person components of X on Y, and their corresponding standard errors. The direction and magnitude of the bias depended on the phase difference between the cycles in the two series.

In cases where cycles existed in only one series, ignoring the cycles typically resulted in biased standard errors of the regression coefficients for the within- and between-person components of X; the direction and magnitude of bias depended on which series contained cyclical components. When cycles only existed in Y, ignoring the cycles led to positive bias in the estimated standard error of the effect of the within-person component of X on Y, but no bias in the estimated standard error of the effect of

the between-person component of X on Y. When cycles only existed in X, ignoring the cycles led to negative bias in the estimated standard error of the effect of the within-person component of X on Y, and negative bias in the estimated standard error of the effect of the between-person component of X on Y.

Composite of Two Effects

When cycles existed in the *X* series and the *Y* series. In Study 1, there were 16 conditions in which the standardized bias of $\hat{\gamma}_{30}$ was greater than .40. Among these 16 conditions, 9 had a standardized bias of $\hat{\sigma}^2$ that was greater than .30 (2 additional conditions had a standardized bias of $\hat{\sigma}^2$ that was greater than .25), even though the ANOVA results did not show an impact of model fitting strategy on the simple bias of $\hat{\sigma}^2$, according to my criterion that η^2 >.01 indicated an appreciable effect. These 9 (or 11) conditions were the only conditions in Study 1 in which the standardized bias of $\hat{\sigma}^2$ was greater than .30 (or .25). In Study 2, there were 6 conditions in which the standardized bias of $\hat{\gamma}_{30}$ was greater than .40, and these were the same 6 conditions in Study 2 in which the standardized bias of $\hat{\sigma}^2$ was greater than .40. All these conditions were conditions in which the existing cycles were ignored. There seems to be a correspondence between the bias in $\hat{\gamma}_{30}$ and the bias in $\hat{\sigma}^2$ resulting from not modeling the cycles. Failure to consider the cyclical components in both X and Y seems to result in biased estimates of the within-person effect of X and overestimated Level-1 residual variance.

In the present studies, when cycles existed in the X series but were ignored, it can be shown by rearranging terms in Equation (19) that the sample estimate of the within-person component of X obtained by person-mean centering is $xc_{it} = \hat{x}_{it} - \bar{x}_i = \hat{e}_{it} +$

$\hat{\mu}_{10}\sin(\frac{2\pi t}{7}) + \hat{\mu}_{20}\cos(\frac{2\pi t}{7})$, where $\hat{e}_{it}$ is the empirical Bayes estimate of the Level-1 residual in the X series. That is to say, with cycles existing in the X series, person-mean centering resulted in an estimate of xw_{it} that contains cyclical components. When cycles also existed in Y but were not modeled, $\hat{\gamma}_{30}$ became a complex composite of a) the regression coefficient of Y on $\hat{e}_{it}$, the empirical Bayes estimate of the Level-1 residual in the X series, and b) the regression of Y on $\hat{\mu}_{10}\sin(\frac{2\pi t}{7}) + \hat{\mu}_{20}\cos(\frac{2\pi t}{7})$, the cyclical components in xc_{it}, the sample estimate of the within-person component of X obtained by person-mean centering. When cycles in the two series were ignored, $\hat{\gamma}_{30}$ should be influenced by the magnitudes of its two components. Hence, it is not surprising that this composite is different from the true value of γ_{30}, resulting in biased $\hat{\gamma}_{30}$. Moreover, the phase difference between the cyclical components in Y and the cyclical components in xc_{it} (which is the same as the cyclical components in the X series), should influence the regression coefficient of Y on $\hat{\mu}_{10}\sin(\frac{2\pi t}{7}) + \hat{\mu}_{20}\cos(\frac{2\pi t}{7})$, and hence influence $\hat{\gamma}_{30}$. Therefore in Studies 1 and 2 where cycles existed not only in the X series but also in the Y series, an interaction was found between model fitting strategy and synchronization of the cycles in the two series (Study 1) or the phase difference between the cycles in the two series (Study 2).

However, due to the different magnitudes of the regression of Y on $\hat{e}_{it}$ and the regression of Y on $\hat{\mu}_{10}\sin(\frac{2\pi t}{7}) + \hat{\mu}_{20}\cos(\frac{2\pi t}{7})$, the cyclical components in Y would not be completely accounted for by xc_{it}, the person-mean centered X, especially when the cycles in the two series were not synchronized. In such cases, it is likely that the unaccounted-for cyclical components in Y would leak into the Level-1 residuals in the

multilevel model of Y, resulting in overestimated $\hat{\sigma}^2$. This in turn could lead to less efficient standard errors at Level-1 (e.g., $\widehat{SE}_{\gamma_{30}}$), and so the statistical tests of the corresponding Level-1 regression coefficients may have lower power. In Study 1 the majority of the conditions with substantial standardized bias of $\hat{\gamma}_{30}$ and $\hat{\sigma}^2$ had cycles that were out of phase by 180 °in the two series, whereas in Study 2, the majority of the conditions with substantial standardized bias of $\hat{\gamma}_{30}$ and $\hat{\sigma}^2$ had cycles in the two series with a phase difference of 3 and 4 days (~180 °).

In Study 2 where the profile of cycles was fixed to having both sine and cosine waves, the ANOVA analysis revealed a material main effect of model fitting strategy on the simple bias in $\hat{\sigma}^2$, where $\hat{\sigma}^2$ was found to be overestimated when the existing cycles were ignored. This main effect was *not* found in Study 1, probably due to the small effect sizes of γ_{10} and γ_{20} in two thirds of the conditions (where the profile of cycles was sine wave only or cosine wave only). Given the small magnitude of the cyclical components in Y that were not accounted for, it is not surprising that the increase in variance of the Level-1 residuals was correspondingly small.

When cycles existed in only the *X* series or only the *Y* series. In Study 3 when cycles only existed in Y but not in X, failure to model the cycles did *not* lead to bias in $\hat{\gamma}_{30}$ (as estimates of the within-person component of X would not contain cyclical components). In such cases, the unaccounted for cyclical components in Y leaked into the Level-1 residuals in the multilevel model of Y, resulting in overestimated $\hat{\sigma}^2$. Although the ANOVA analysis did not reveal an η^2 value greater than .01, $StdB_{\hat{\sigma}^2}$ was .407 when the Level-1 residuals in Y were independent in the population model, but was .158 when the Level-1 residuals in Y were Lag-1 auto correlated in the population model.

In Study 3 when cycles only existed in X but not in Y, not modeling the cycles turned $\hat{\gamma}_{30}$ into a complex composite of a) the regression coefficient of Y on $\hat{e}_{it}$ (whose expected value should be the true value of γ_{30}, .5477), and b) the regression of Y on $\hat{\mu}_{10}\sin(\frac{2\pi t}{7}) + \hat{\mu}_{20}\cos(\frac{2\pi t}{7})$, which should be 0 since by design cycles did not existed in Y. In such cases, even though the ANOVA analysis did not reveal an η^2 value greater than .01, $\hat{\gamma}_{30}$ was underestimated ($StdB_{\hat{\gamma}_{30}}$ was -.371 when the Level-1 residuals in Y were independent in the population model, but was -.236 when the Level-1 residuals in Y were Lag-1 auto correlated in the population model).

Bias in Estimation of the Fixed Effect Representing the Level-2 X-Y Relationship

For $\hat{\gamma}_{02}$, the estimated fixed effect of the between-person component of X on Y, Study 1 found a trend for an interaction between Model Fitting Strategy and Magnitude of Standardized Regression Coefficient for xw_{it} ($\eta^2 = .005$). Modeling the cycles in the two series always resulted in unbiased $\hat{\gamma}_{02}$ (mean $StdB_{\hat{\gamma}_{02}} = .014$, .012, and .003 when the standardized regression coefficient for xw_{it} was 0, .20, and .50, respectively). In contrast, not modeling the cycles tended to lead to biased $\hat{\gamma}_{02}$, with the direction and magnitude of bias depending on the Magnitude of Standardized Regression Coefficient for xw_{it} (mean $StdB_{\hat{\gamma}_{02}} = -.125$, .008, and .204 when the standardized regression coefficient for xw_{it} was 0, .20, and .50, respectively). In Study 2 where the Magnitude of Standardized Regression Coefficient for xw_{it} was fixed to .50 and Profile of Cycles was fixed to both sine and cosine waves, an impact of Model Fitting Strategy ($\eta^2 = .011$) was found. Ignoring the existing cycles in the two series resulted in positively biased $\hat{\gamma}_{02}$ (mean $StdB_{\hat{\gamma}_{02}} = .204$), whereas modeling the cycles resulted in unbiased $\hat{\gamma}_{02}$ (mean

$StdB_{\hat{\gamma}_{02}} = .002$). From the results of Studies 1 and 2, I conjecture that the mechanism is related to the Magnitude of Standardized Regression Coefficient for xw_{it}.

In Study 3, I found no effect of model fitting strategies. However, a closer examination revealed that when cycles only existed in Y and when the Level-1 residuals in Y were independent, $\hat{\gamma}_{02}$ was positively biased *whether or not* the cycles were modeled ($StdB_{\hat{\gamma}_{02}} = .255$ in both cases). When the Level-1 residuals in Y were Lag-1 autocorrelated, whether or not the cycles that only existed in Y were modeled, $\hat{\gamma}_{02}$ was slightly positively biased ($StdB_{\hat{\gamma}_{02}} = .110$ in both cases). When cycles only existed in X, however, when the Level-1 residuals in Y were independent, $\hat{\gamma}_{02}$ was positively biased when the cycles were ignored ($StdB_{\hat{\gamma}_{02}} = .276$), but rather close to its true parameter value when the cycles were modeled ($StdB_{\hat{\gamma}_{02}} = .063$). When the Level-1 residuals in Y were Lag-1 autocorrelated, this pattern remained, but the magnitude of bias in $\hat{\gamma}_{02}$ became smaller ($StdB_{\hat{\gamma}_{02}} = .181$ when the cycles were ignored, and $StdB_{\hat{\gamma}_{02}} = .002$ when the cycles were modeled). The conditions in Study 3 when cycles only existed in Y were the only conditions in the three Monte Carlo studies in which I used the person-mean centering strategy to disaggregate the within- and between-person components of X whether or not the cycles were modeled. In other conditions in Study 3 and in Studies 1 and 2, the person-mean centering strategy was used only when the cycles were ignored, whereas when the cycles were modeled, the empirical Bayes estimates of the Level-1 residuals and the Level-2 intercept residuals were taken as estimates of the within- and between-person components of X. Given (a) the strategies used to disaggregate the within- and between-person components of X and (b) the patterns of $StdB_{\hat{\gamma}_{02}}$ in the 3

studies, I conjecture that the mechanism might be related to unreliability of the person mean in representing the between-person component of X (Lüdtke, Marsh, Robitzsch, Trautwein, Asparouhov, & Muthén, 2008).

Lüdtke et al. (2008) demonstrated using mathematical derivations and simulations that using person means as the between-person component of a Level-1 predictor X can result in biased estimates of the effect of X at Level-2 in the two level multilevel model of Y, when the model of Y is correctly specified. They provided an equation (Equation (7) in Lüdtke et al., 2008) for the expected bias in the estimates of the regression coefficient of X at Level-2, reproduced and adapted to the setting in the current studies as

$$E(\hat{\gamma}_{between} - \gamma_{between}) = (\gamma_{within} - \gamma_{between}) \cdot \frac{1}{n} \cdot \frac{(1-ICC)}{ICC+(1-ICC)/n}. \quad (26)$$

ICC is the intraclass correlation coefficient of the Level-1 predictor X, n is the number of observations within each Level-2 unit, and γ_{within} and $\gamma_{between}$ are the true values of the Level-1 and Level-2 unstandardized regression coefficients for the X-Y relationship. Lüdtke et al. (2008) did not consider potential effects like cycles since they focused on the classic multilevel context in which Level-1 represents persons and Level-2 represents groups. Hence, their mathematically derived equation for the expected bias in the estimates of the effect of X at Level-2 cannot be directly applied to the present studies. To get a better idea of the influence of the unreliability of the person means of X, consider a situation in which there are no cycles in the population model for either X or Y in the present studies.

In the present studies the number of observations within each Level-2 unit was held constant at 63. Since the magnitude of γ_{02} was held constant at .2236, the difference between the Level-1 and Level-2 regression coefficients for the X-Y relationship was -.2236, .0045, and .3241 when the Magnitude of Standardized Regression Coefficient for xw_{it} was 0, .20, and .50[15], respectively. If there were no cycles in the population model of both X and Y, the ICC of X would be .167[16]. Hence, when the Magnitude of Standardized Regression Coefficient for xw_{it} was 0, .20, and .50, the expected raw bias in $\hat{\gamma}_{02}$ due to using the person mean as the between-person component of X should be -.016, .003, and .024, respectively. These values are very close to the mean raw bias values observed in the present studies. In Study 1, when the cycles were not modeled (hence the person mean was used as the between-person component of X), the mean raw bias in $\hat{\gamma}_{02}$ was -.015, .001, and .025, respectively (corresponding to mean $StdB_{\hat{\gamma}_{02}}$ values of -.125 .008, and .204). In Study 2, when the cycles were not modeled (hence the person mean was used as the between-person component of X), the mean raw bias in $\hat{\gamma}_{02}$ was.025 (corresponding to mean $StdB_{\hat{\gamma}_{02}}$ value of .204). There is obvious similarity between the mean raw bias in $\hat{\gamma}_{02}$ observed in Studies 1 and 2, and the expected raw bias in $\hat{\gamma}_{02}$ due to using the person mean as the between-person component of X derived from population models of X and Y that have no cycles but everything else the same as the settings in Studies 1 and 2. This indicates that, at least in the settings of Studies 1 and 2

[15] These correspond to unstandardized γ_{30} values of 0, .2191, and .5477, respectively.
[16] This is not very much different from the ICC of X in the present studies. Given the small magnitudes of cyclical components in the present studies, the influence of them on the ICC of X was in the third decimal place.

where cycles existed in both series, the bias in $\hat{\gamma}_{02}$ mainly comes from unreliability of the person mean as an estimate of the between-person component of X.

In Lüdtke et al. (2008), the Level-1 residuals in Y were always independent. Importantly, in my Studies 1 and 2 whether the Level-1 residuals in Y were independent or Lag-1 autocorrelated did not make any difference in terms of bias in $\hat{\gamma}_{02}$. In Study 3, however, the independence of the residuals seems to have some impact. When cycles only existed in Y and when the Level-1 residuals in Y were specified as independent (the person mean was used as the between-person component of X *whether or not* the cycles were modeled), the mean raw bias in $\hat{\gamma}_{02}$ was .025 in both cases ($StdB_{\hat{\gamma}_{02}} = .255$ in both cases). When the Level-1 residuals in Y were Lag-1 autocorrelated (the person mean was used as the between-person component of X *whether or not* the cycles were modeled), the mean raw bias in $\hat{\gamma}_{02}$ was .014 in both cases ($StdB_{\hat{\gamma}_{02}} = .110$ in both cases). When cycles only existed in X and when the Level-1 residuals in Y were independent, when the cycles were not modeled (hence the person mean was used as the between-person component of X), the mean raw bias in $\hat{\gamma}_{02}$ was .031 ($StdB_{\hat{\gamma}_{02}} = .276$). When the Level-1 residuals in Y were Lag-1 autocorrelated, when the cycles were not modeled (hence the person mean was used as the between-person component of X), the mean raw bias in $\hat{\gamma}_{02}$ was .025 ($StdB_{\hat{\gamma}_{02}} = .181$). There is small deviation between the mean raw bias in $\hat{\gamma}_{02}$ observed in Study 3, and the expected raw bias in $\hat{\gamma}_{02}$ due to the use of the person mean as the between-person component of X derived from population models of X and Y that have no cycles but otherwise have identical parameters to those in Study 3. This pattern of results indicates that at least in the settings of Study 3 where cycles existed in only one

of the series, the series contains cyclical components and the Level-1 residual variance-covariance structure in Y potentially can influence the bias in $\hat{\gamma}_{02}$. Results of Studies 1 and 2 indicate that unreliability of the person mean as an estimate of the between-person component of X could lead to bias in $\hat{\gamma}_{02}$. Results from Study 3 seem to suggest that this impact of unreliability in the estimate of the between-person component of X is modified by which cycles contains cycles and the Level-1 residual variance-covariance structure in Y.

Bias in Estimation of the Standard Errors of the Fixed Effect Representing the Level-2 *X*-*Y* Relationship

In Studies 1 and 2, and in Study 3 when the cycles only existed in the X series, not modeling the existing cycles led to similar underestimates of $\widehat{SE}_{\gamma_{02}}$ as compared to when the cycles were modeled: the magnitude of the mean RB of $\widehat{SE}_{\gamma_{02}}$ in these conditions were essentially the same (-.076 to -.077). That is to say, in Study 1 when the unstandardized regression coefficients representing the Level-1 and Level-2 X- Y relationships were similar in magnitude, not modeling the cycles (or rather, using the person mean as an estimate of the between-person component of X) did not lead to bias in $\hat{\gamma}_{02}$, but rather led to bias in $\widehat{SE}_{\gamma_{02}}$. This might make the bias in $\widehat{SE}_{\gamma_{02}}$ worthy of attention.

The sampling variances (squared standard errors) of regression coefficients in the multilevel modeling framework are the diagonal elements of $\Sigma_{\hat{\gamma}} = (\sum_{i=1}^{100} \boldsymbol{X}_i^T \boldsymbol{V}_i^{-1} \boldsymbol{X}_i)^{-1}$, where $\boldsymbol{V}_i = \boldsymbol{Z}_i \tau_{00} \boldsymbol{Z}_i^T + \sigma^2 \boldsymbol{I}_{63}$ in the present studies. $\boldsymbol{X}_i$ is the 63 $\times k$ design matrix containing the values of the intercept and the (k-1) predictors for person i in the 63 time points. $\boldsymbol{Z}_i$ is the 63 $\times u$ design matrix containing the values of predictors with random

effects (u=1: intercept). $\boldsymbol{I}_{63}$ is a 63 × 63 identity matrix. When cycles were not modeled, $\boldsymbol{Z}_i$ which contains a column of 1's was not changed, $\hat{\tau}_{00}$ was not biased, $\hat{\sigma}^2$ was overestimated in some but not all conditions. Hence, the mechanism that led to underestimation of $\widehat{SE}_{\gamma_{02}}$ when the cycles (in the X series) were ignored appears to be unrelated to $\boldsymbol{V}_i$. The mechanism should be related to the change in the design matrix $\boldsymbol{X}_i$ when cycles were *not* modeled versus when cycles were modeled. There were 3 changes in the design matrix $\boldsymbol{X}_i$ in Studies 1 and 2: a) the cyclical components in Y were in the design matrix $\boldsymbol{X}_i$ when cycles were modeled, but not when the cycles were ignored; b) the predictor variable $\widehat{xw}_{it}$ in the design matrix $\boldsymbol{X}_i$ was the empirical Bayes estimate of the Level-1 residual in the X series (which did not include cyclical components) when cycles were modeled, but was xc_{it}, the person-mean centered X (which included cyclical components) when cycles were *not* modeled; and c) the predictor variable $\widehat{xb}_i$ in the design matrix $\boldsymbol{X}_i$ was the empirical Bayes estimate of the Level-2 intercept residual in the X series ($= \bar{x}_i - \hat{\mu}_{00}$) when cycles were modeled, but was $\bar{x}_i$, the person-mean of X when cycles were *not* modeled. In Study 3, when cycles only existed in the Y series, not modeling cycles resulted in unbiased $\widehat{SE}_{\gamma_{02}}$; in contrast, when cycles only existed in the X series, not modeling cycles resulted in underestimated $\widehat{SE}_{\gamma_{02}}$. When cycles only existed in the X series, there were no cyclical components in Y in the design matrix $\boldsymbol{X}_i$ whether or not the cycles were modeled, so that change a) in the design matrix $\boldsymbol{X}_i$ should not be the cause of underestimated $\widehat{SE}_{\gamma_{02}}$. This initial conclusion implies that the cause should be change b) or change c) in the design matrix $\boldsymbol{X}_i$. However, if change b) were the cause of underestimated $\widehat{SE}_{\gamma_{02}}$, in Study 1 the magnitude of the cyclical components in the X

series (and thus person-mean centered X) should have an impact on the magnitude of RB of $\widehat{SE}_{\gamma_{02}}$. However, this result was not found in Study 1. If change c) were the cause of underestimated $\widehat{SE}_{\gamma_{02}}$, this conclusion would be consistent with the finding in Lüdtke et al. (2008) that using the person mean as an estimate of the between-person component of X can substantially underestimate the associated standard error. Moreover, in Study 3 when the cycles only existed in the Y series, the person-mean centering strategy was used to disaggregate the within- and between-person components of X, whether or not cycles in Y were modeled. Hence, if change c) were the cause of underestimated $\widehat{SE}_{\gamma_{02}}$ in Studies 1 and 2, in Study 3 not modeling the cycles in the Y series should not lead to smaller $\widehat{SE}_{\gamma_{02}}$ as compared to when the cycles in Y were modeled. This inference is consistent with the obtained findings in Study 3. Therefore, using the person mean as an estimate of the between-person component of X appears to lead to substantial underestimation of the associated standard error as compared to using the empirical Bayes estimates of the Level-2 intercept residuals in X.

Some Limitations of the Present Studies

The effect sizes for the cyclical effects used in the present studies characterized those of Beal and Ghandour (2011), a substantive daily diary data study that used the sine and cosine modeling approach to model the weekly cycle and reported the magnitudes of the weekly cycle. However, these are *very small* effect sizes in Cohen's (1988) terms, and thus some important effects may lack sufficient variability to be detected by the ANOVA analyses. For example, there were a few interaction effects on the simple bias of $\hat{\gamma}_{30}$ involving serial dependency in the outcome series or the profile of cycles with an η^2

value smaller than .01 but greater than .005. There were a few main effects of (or interaction effect involving) the model fitting strategy on $\hat{\sigma}^2$ with an η^2 value smaller than .01 but greater than .005. These potentially important effects may have been masked by random error in the present studies. These effects should be investigated using greater effect sizes for the cyclical effects in the future.

Another limitation of the present studies is that the combined sine and cosine functions cannot model all possible forms of weekly cycles. For example, sine and cosine functions cannot model the weekly cycles in the number of drinks consumed and the desire to drink in Armeli et al. (2000) where the highest positive value at the end of one cycle was followed by a negative value at the very beginning of the next cycle. Profiles of cycles that do not follow the sine and cosine functions could be considered in future simulations.

REFERENCES

Armeli, S., Carney, M. A., Tennen, H., Affleck, G., & O'Neil, T. P. (2000). Stress and alcohol use: A daily process examination of the stressor-vulnerability model. *Journal of Personality and Social Psychology*, *78*(5), 979-994.

Beal, D. J. & Ghandour, L. (2011). Stability, change, and the stability of change in daily workplace affect. *Journal of Organizational Behavior*, *32*(4), 526-546.

Biesanz, J. C., West, S. G., & Kwok, O. (2003). Personality over time: Methodological approaches to the study of short-term and long-term development and change. *Journal of Personality*, *71*(6), 905-941.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*(1), 579-616.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Mahwah, NJ: Erlbaum.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330-351.

Congdon, P. (2006). *Bayesian Statistical Modeling* (2nd Ed.). New York: John Wiley & Sons.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121-138.

Enders, C. K., & Tofighi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling*, *15*(1), 75-95.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods*, *9*(4), 466-491.

Gable, S. L., & Reis, H. T., (1999). Now and then, them and us, this and that: Studying relationships across time, partner, context, and person. *Personal Relationships*, *6*(4), 415-432.

Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivariate Behavioral Research, 47*(1), 1-25.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.

Kreft, I. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*(1), 1-21.

Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review, 23*(4), 418-444.

Kwok, O., West, S. G., Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research, 42*(3), 557-592.

Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology*, *58*(1), 164-171.

Larsen, R. J., Augustine, A. A., & Prizmic, Z. (2009). A process approach to emotion and personality: Using time as a facet of data. *Cognition and Emotion*, *23*(7), 1407-1426.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203-229.

McFarlane, J. M., Martin, C. L., & Williams, T. M. (1988). Mood fluctuations: Women versus men and menstrual versus other cycles. *Psychology of Women Quarterly, 12*(2), 201-223.

Murphy, D. L., & Pituch, K. A. (2009). The performance of multilevel growth curve models under an autoregressive moving average process. *Journal of Experimental Education*, *77*(3), 255-282.

Murphy, D. L., Beretvas, S. M., & Pituch, K. A. (2011). The effects of autocorrelation on the curve-of-factors growth model. *Structural Equation Modeling*, *18*(3), 430-448.

Ram, N., Chow, S., Bowles, R. P., Wang, L, Grimm, K., Fujita, F., & Nesselroade, J. R. (2005). Examining interindividual differences in cyclicity of pleasant and

unpleasant affects using spectral analysis and item response modeling. *Psychometrika*, *70*(4), 773-790.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (Second Edition). Thousand Oaks: Sage Publications.

Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods*, *5*(2), 199-213.

Raudenbush, S. W., & Liu, X. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods*, *6*(4), 387-401.

Reis, H. T., & Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. In H. T. Reis & C. M. Judd (Ed.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 190-222). Cambridge: Cambridge University Press.

Ryu, E. West, S. G. & Sousa, K. H. (2012). Between-person and within-person relationships in longitudinal data: Latent intercept model with correlated residuals and multivariate multilevel model. *Annals of Behavioral Medicine*, *43*(3), 330-342.

Velicer, W. F., & Molenaar, P. C. (2012). Time series analysis for psychological research. In I. B. Weiner, J. A. Schinka and W. F. Velicer (Eds.), *Handbook of Psychology*, Volume 2, *Research Methods in Psychology* (2nd Ed., pp. 628-660). Hoboken, NJ: Wiley.

West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality*, *59*(3), 609-662.

West, S. G., Ryu, E., Kwok, O., & Cham, H. (2011). Multilevel modeling: Current and future applications in personality research. *Journal of Personality*, *79*(1), 2-50.

Table 1. Summary of key parameter values in Study 1.

Parameter	Level(s)		
Time-varying predictor series (X):			
$$x_{it} = \mu_{00} + \mu_{10}\sin(\frac{2\pi t}{7}) + \mu_{20}\cos(\frac{2\pi t}{7}) + (\zeta_{0i} + e_{it}).$$			
μ_{00} (Intercept)		0	
μ_{10} (Sine cycle)	0		.0538
μ_{20} (Cosine cycle)	0		.1076
τ_{x00} (Intercept residual variance)		.20	
σ_x^2 (Level-1 residual variance)		1	
Outcome series:			
$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi t}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi t}{7}\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}).$$			
γ_{00} (Intercept)		0	
γ_{10} (Sine cycle)	-.0538	0	.0538
γ_{20} (Cosine cycle)	-.1076	0	.1076
γ_{30} (Effect of xw_{it}, within-person component of X)	0	.2191	.5477
γ_{01} (Effect of z_i, time-invariant predictor, on intercept)		.2236	
γ_{02} (Effect of xb_i, between-person component of X, on intercept)		.2236	
τ_{00} (Intercept residual variance)		.20	
σ^2 (Level-1 residual variance)		1	
ϕ (Autoregressive correlation coefficient)	0		.80

Table 2. Summary of key parameter values in Study 2.

Parameter	Level(s)
Time-varying predictor series (X):	
$$x_{it} = \mu_{00} + \mu_{10}\sin(\frac{2\pi t}{7}) + \mu_{20}\cos(\frac{2\pi t}{7}) + (\zeta_{0i} + e_{it}).$$	
μ_{00} (Intercept)	0
μ_{10} (Sine cycle)	.0538
μ_{20} (Cosine cycle)	.1076
τ_{x00} (Intercept residual variance)	.20
σ_x^2 (Level-1 residual variance)	1
Outcome series:	
$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi(t-PD)}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi(t-PD)}{7}\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}).$$	
γ_{00} (Intercept)	0
γ_{10} (Sine cycle)	.0538
γ_{20} (Cosine cycle)	.1076
γ_{30} (Effect of xw_{it}, within-person component of X)	.5477
γ_{01} (Effect of z_i, time-invariant predictor, on intercept)	.2236
γ_{02} (Effect of xb_i, between-person component of X, on intercept)	.2236
τ_{00} (Intercept residual variance)	.20
σ^2 (Level-1 residual variance)	1
ϕ (Autoregressive correlation coefficient)	0, .80
PD (Phase Difference of the cycles in the two series, in unit of days)	0, 1, 2, 3, 4, 5, 6

Table 3. Summary of key parameter values in Study 3.

Parameter	Level(s)		
Time-varying predictor series (X):			
$$x_{it} = \mu_{00} + \mu_{10}\sin(\frac{2\pi t}{7}) + \mu_{20}\cos(\frac{2\pi t}{7}) + (\zeta_{0i} + e_{it}).$$			
μ_{00} (Intercept)		0	
μ_{10} (Sine cycle)	0		.0538
μ_{20} (Cosine cycle)	0		.1076
τ_{x00} (Intercept residual variance)		.20	
σ_x^2 (Level-1 residual variance)		1	
Outcome series:			
$$y_{it} = \gamma_{00} + \gamma_{10}\sin\left(\frac{2\pi t}{7}\right) + \gamma_{20}\cos\left(\frac{2\pi t}{7}\right) + \gamma_{30}xw_{it} + \gamma_{01}z_i + \gamma_{02}xb_i + (\xi_{0i} + \varepsilon_{it}).$$			
γ_{00} (Intercept)		0	
γ_{10} (Sine cycle)	0		.0538
γ_{20} (Cosine cycle)	0		.1076
γ_{30} (Effect of xw_{it}, within-person component of X)		.5477	
γ_{01} (Effect of z_i, time-invariant predictor, on intercept)		.2236	
γ_{02} (Effect of xb_i, between-person component of X, on intercept)		.2236	
τ_{00} (Intercept residual variance)		.20	
σ^2 (Level-1 residual variance)		1	
ϕ (Autoregressive correlation coefficient)	0		.80

Table 4. Comparison of key parameters estimated using the sandwich estimator (with DDFM = BETWITHIN) versus using the model-based estimator (with DDFM = KenwardRoger)

Condition	Model fitting strategy	Estimator	$\hat{\gamma}_{30}$			$\hat{\gamma}_{02}$		
			Std. bias	$\overline{RB_{\widehat{SE}_{\gamma_{30}}}}$	Coverage of 95% CI's	Std. bias	$\overline{RB_{\widehat{SE}_{\gamma_{02}}}}$	Coverage of 95% CI's
1	Modeling cycles	Model-based	0.019	0	0.941	0.048	0	0.945
	Modeling cycles	Sandwich	0.019	-0.007	0.934	0.048	-0.014	0.936
	Not modeling cycles	Model-based	-0.839	0.005	0.85	0.268	-0.077	0.933
	Not modeling cycles	Sandwich	-0.839	-0.002	0.84	0.268	-0.089	0.928
2	Modeling cycles	Model-based	0.015	0	0.956	0.047	0	0.932
	Not modeling cycles	Model-based	0.057	-0.001	0.95	0.259	-0.077	0.924
	Not modeling cycles	Sandwich	0.057	-0.009	0.95	0.259	-0.095	0.917
3	Modeling cycles	Model-based	-0.007	0	0.938	0.11	0	0.943
	Not modeling cycles	Model-based	-0.009	0.007	0.942	0.11	0	0.943
	Not modeling cycles	Sandwich	-0.009	-0.005	0.93	0.11	-0.015	0.937

Note. In Condition 1, the cycles had sine and cosine waves, cycles in the X and Y series were out of phase by 180 °, the Level-1 residuals in Y were independent, and the standardized regression coefficient for xw_{it} was .50. In Condition 2, the cycles had sine and cosine waves, the cycles in Y lagged 1 day behind the cycles in X, the Level-1 residuals in Y were independent, and the standardized regression coefficient for xw_{it} was .50. In Condition 3, cycles only existed in the outcome series and had sine and cosine waves, the Level-1 residuals in Y were Lag-1 autocorrelated, and the standardized regression coefficient for xw_{it} was .50.

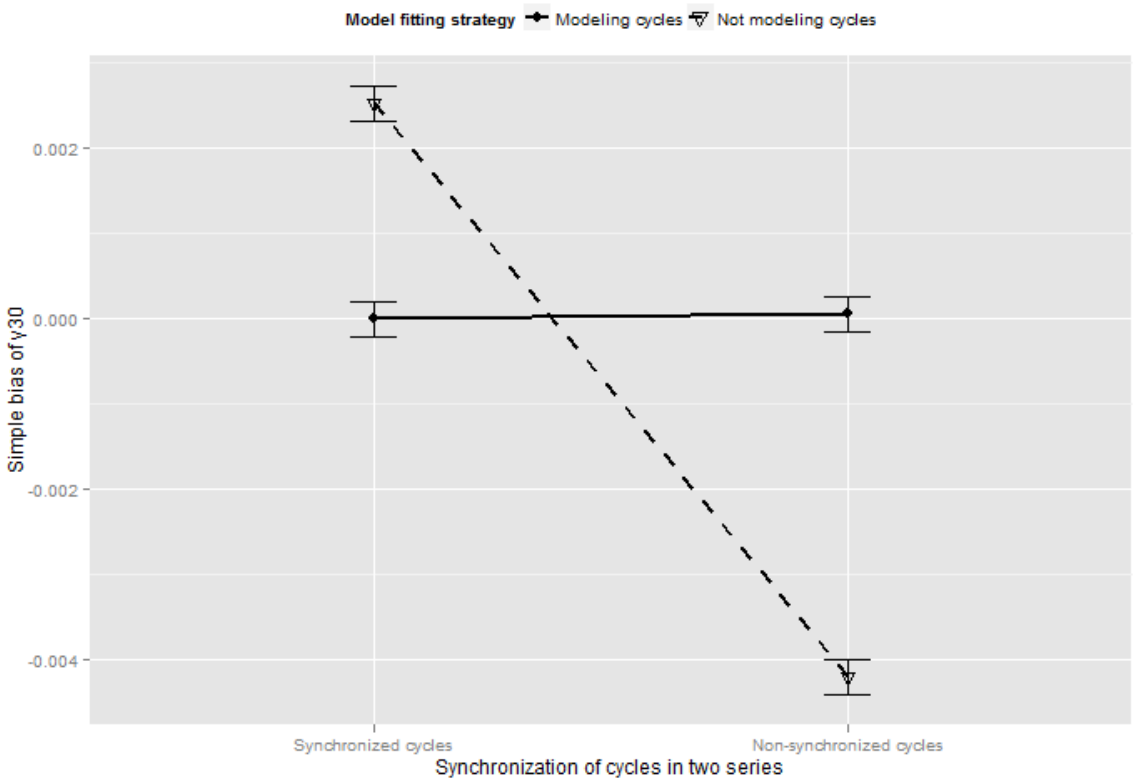

Figure 1. Simple bias of $\hat{\gamma}_{30}$ by Model Fitting Strategy and Synchronization, averaged over the remaining design cells. The error bars represent the 95% confidence intervals.

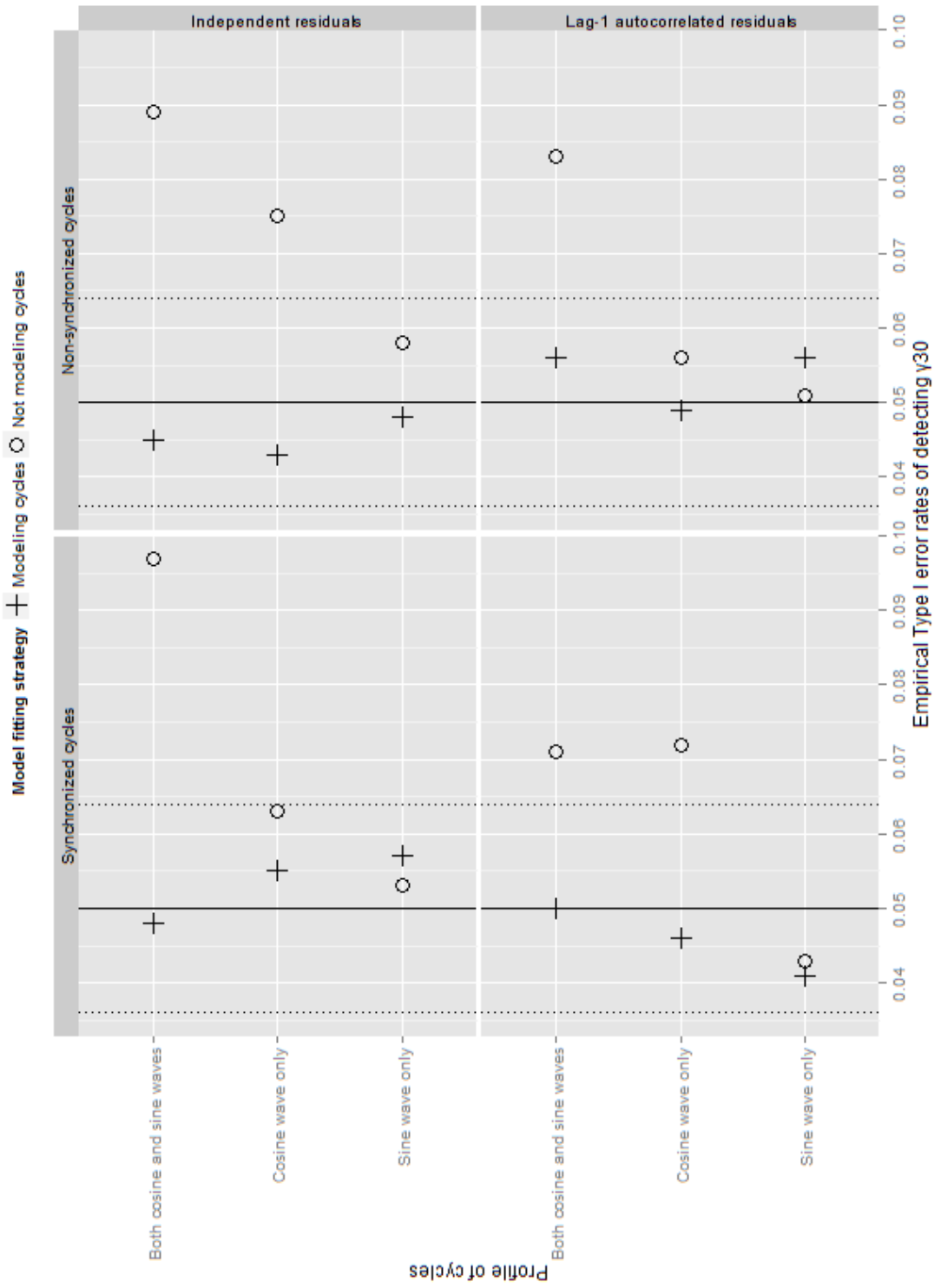

Figure 2. Empirical Type I error rates of detecting $\hat{\gamma}_{30}$. Dotted lines indicate the range for acceptable Type I error rates.

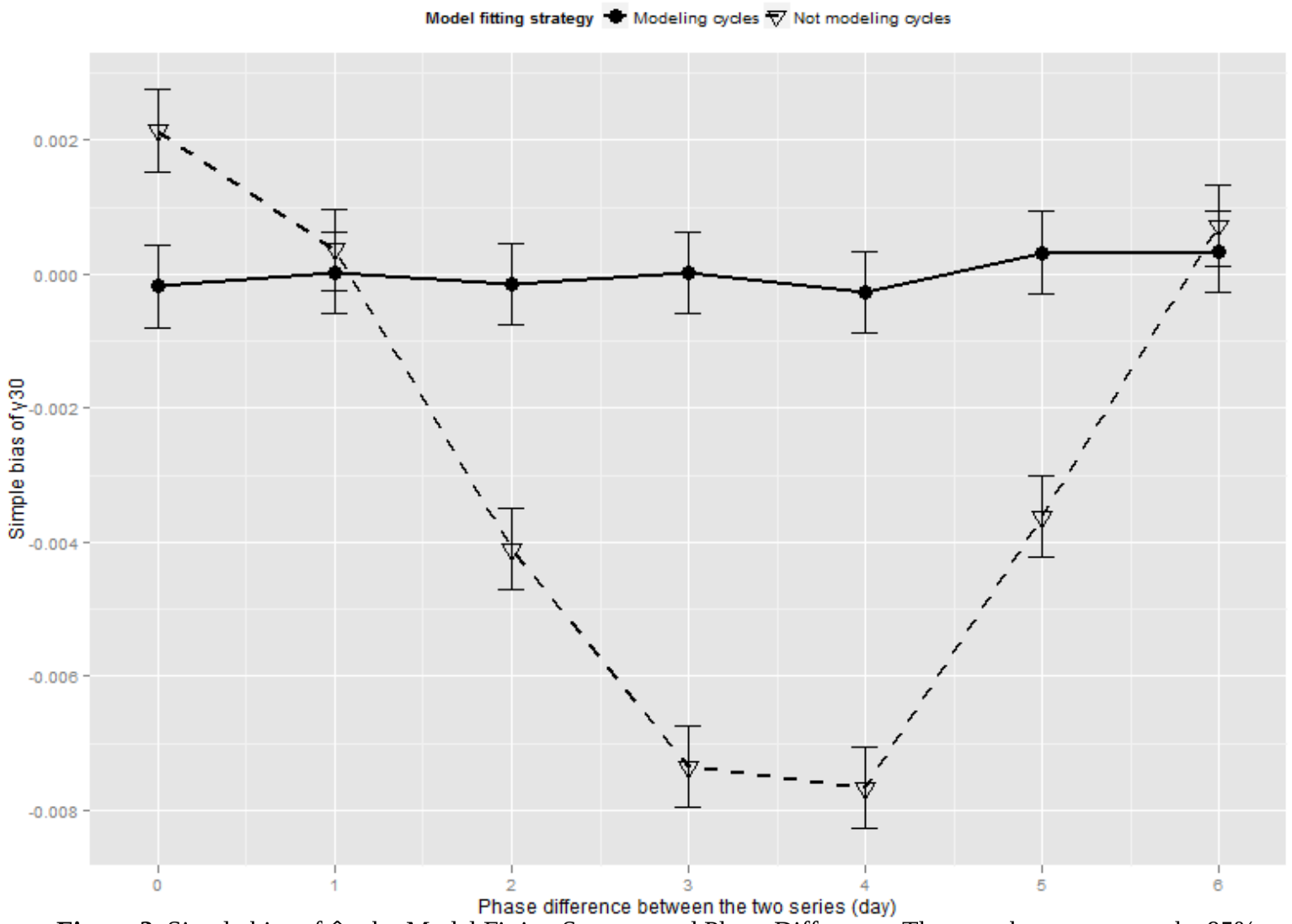

Figure 3. Simple bias of $\hat{\gamma}_{30}$ by Model Fitting Strategy and Phase Difference. The error bars represent the 95% confidence intervals.

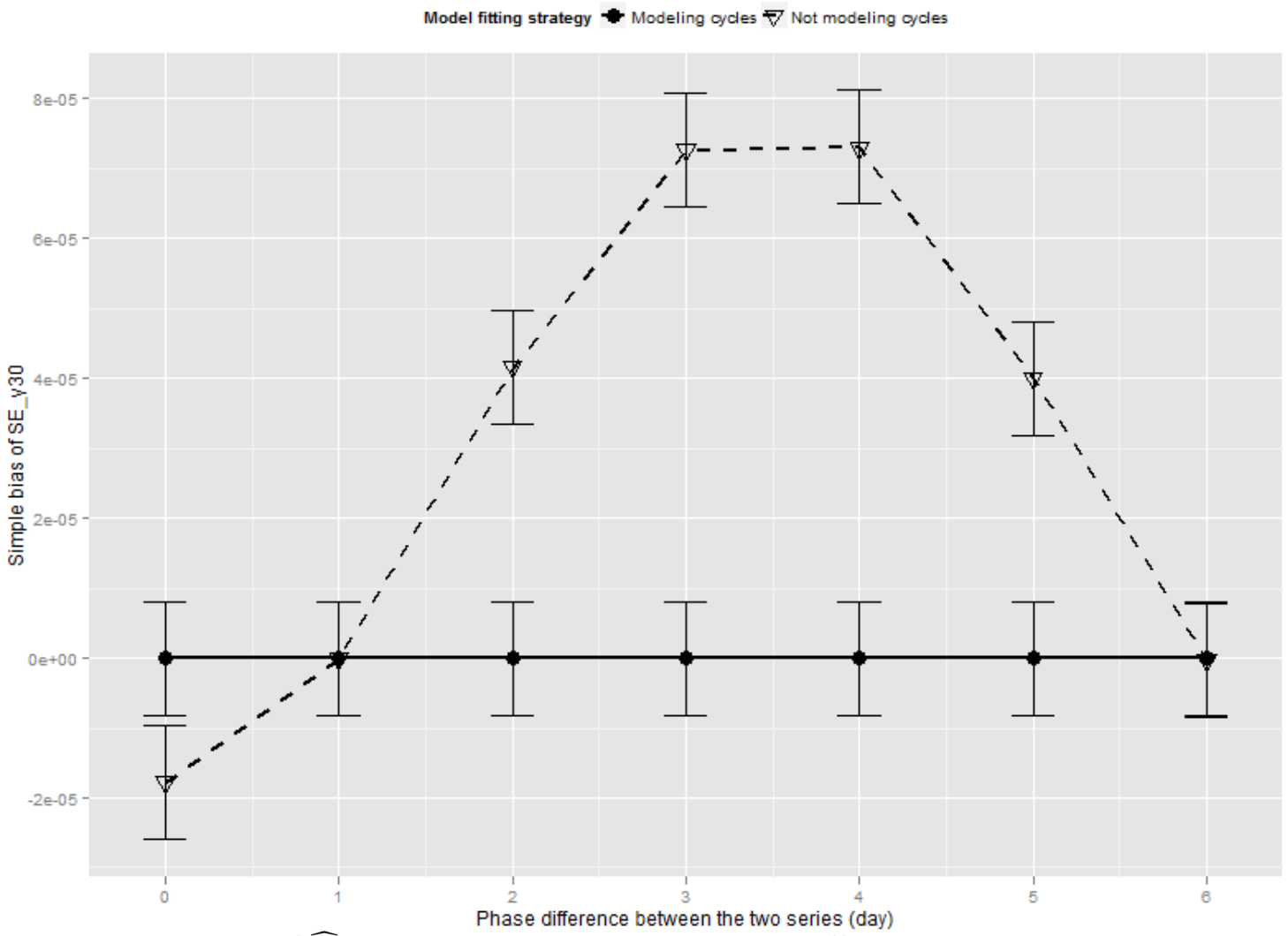

Figure 4. Simple bias of $\widehat{SE}_{\gamma_{30}}$ by Model Fitting Strategy and Phase Difference, averaged over the remaining design cells. The error bars represent the 95% confidence intervals.

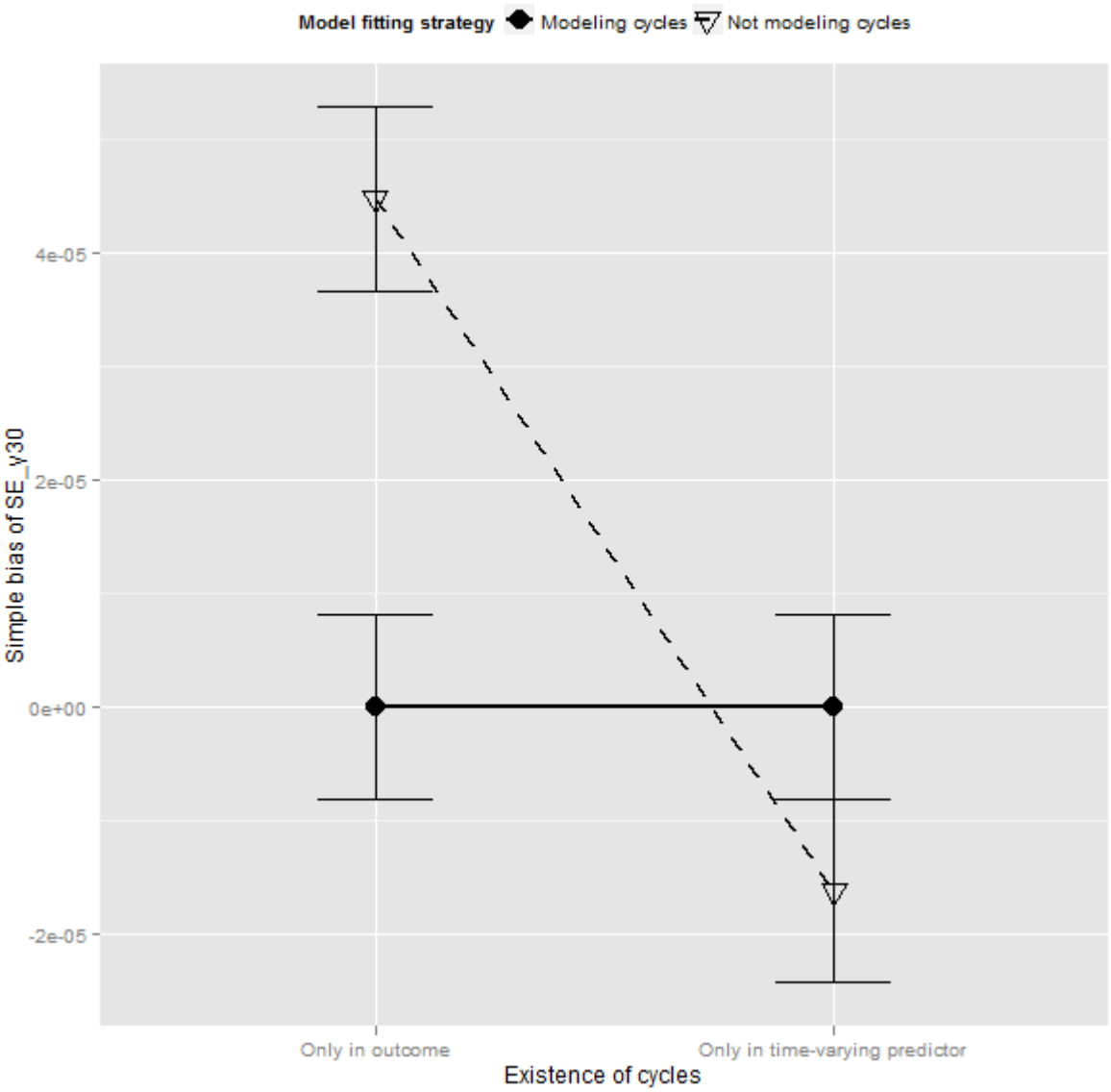

Figure 5. Simple bias of $\widehat{SE}_{\gamma_{30}}$ by Model Fitting Strategy and Existence of Cycles, averaged over the remaining design cells. The error bars represent the 95% confidence intervals.

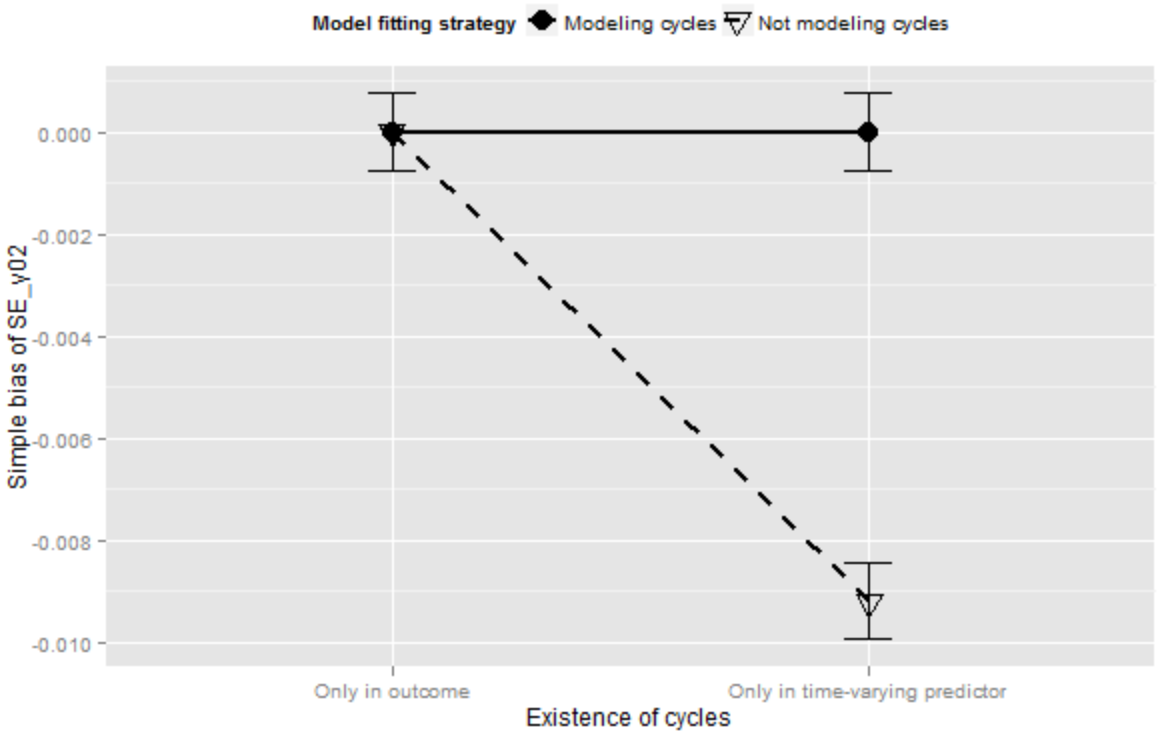

Figure 6. Simple bias of $\widehat{SE}_{\gamma_{02}}$ by Model Fitting Strategy and Existence of Cycles, averaged over the remaining design cells. The error bars represent the 95% confidence intervals.

APPENDIX A

SAS SCRIPT FOR DATA GENERATION

a) Study 1 script for conditions in which the Level-1 residuals in the outcome series were independent

```
/* gamma_X00  -- mean intercept in the X series                */
/* gamma_X10  -- fixed effect for sine function in the X series       */
/* gamma_X20  -- fixed effect for cosine function in the X series */
/* gamma00    -- mean intercept in the Y series                                */
/* gamma10    -- fixed effect for sine function in the Y series          */
/* gamma20    -- fixed effect for cosine function in the Y series */
/* gamma30    -- fixed effect for Xw_it in the Y series                  */
/* gamma01    -- mean difference in intercept of Y caused by TICi  */
/* gamma02    -- mean difference in intercept of Y caused by Xb_i */
/* AR1    -- Level-1 residual autoregressive correlation     */

data sample;
* (Outer Loop) Sample the level-2 units: person;
  do person=1 to &nperson;  *<-- nperson is the number of individuals
        * Random effects in X, the time-varying predictor;
    rx1=rannor(-1);
    zeta0i=rx1*sqrt(&taux00);  *<-- zeta0i is random intercept in X, Variance = taux00;
        * Random effects in Y, the outcome;
    r1=rannor(-1);
    xi_0i=r1*sqrt(&tau00);   *<-- xi_0i is random intercept in Y, Variance = tau00;
```

```
        * TIC: time invariant preditor;
      TICi=rannor(-1);


* (Inner loop) Within a level-2 unit, sample the level-1 units: time;
        do time=&st_time to &end_time by 1;
    * st_time and end_time are the first and last measurement occasions;
                sin_Func=sin((2*CONSTANT('PI')*time)/7);
                cos_Func=cos((2*CONSTANT('PI')*time)/7);
                * for X;
                eit=rannor(-1);  *<-- Level-1 residual in X, Variance equals 1;
                Xit=&gamma_X00
                        + &gamma_X10*sin_Func + &gamma_X20*cos_Func
                        + zeta0i + eit;
                Xb_i=zeta0i;
                Xw_it=eit;
                * for Y;
                epsilon_it=rannor(-1);
                Yit=&gamma00 + &gamma10*sin_Func + &gamma20*cos_Func
                        + &gamma30*Xw_it
                        + &gamma01*TICi + &gamma02*Xb_i
                        + xi_0i + epsilon_it;
        output;
        end;  * <-- end inner loop, go to next level-1 unit;
```

```
   end;  * <-- end outer loop, go to next level-2 unit;
run;
```

b) Study 1 script for conditions in which the Level-1 residuals in the outcome series were Lag-1 autocorrelated

```
data sample;
* (Outer Loop) Sample the level-2 units: person;
   do person=1 to &nperson;
         * Random effects in X, the time-varying predictor;
      rx1=rannor(-1);
      zeta0i=rx1*sqrt(&taux00);  *<-- random intercept in X, Variance = taux00;
         * Random effects in Y, the outcome;
      r1=rannor(-1);
      xi_0i=r1*sqrt(&tau00);   *<-- random intercept in Y, Variance = tau00;
         * TIC: time invariant preditor;
      TICi=rannor(-1);

* (Inner loop) Within a level-2 unit, sample the level-1 units: time;
         epsilon_it=rannor(-1); *epsilon_i(-31);
         do time=&st_time to &end_time by 1;
               sin_Func=sin((2*CONSTANT('PI')*time)/7);
               cos_Func=cos((2*CONSTANT('PI')*time)/7);
```

```
   * for X;
   eit=rannor(-1);  *<-- Level-1 residual in X, Variance equals 1;
   Xit=&gamma_X00 + &gamma_X10*sin_Func + &gamma_X20*cos_Func
            + zeta0i + eit;
   Xb_i=zeta0i;
   Xw_it=eit;
   * for Y;
   if time = &st_time then do; * t=-31;
     epsilon_last=0;
     u_it=0;
   end;
   else do; * t>-31;
* value of epsilon from the previous time point;
     epsilon_last=epsilon_it;
     u_it=rannor(-1)*sqrt(1-&AR1**2);
* value of epsilon at the current time point;
     epsilon_it=&AR1*epsilon_last + u_it;
   end;
   Yit=&gamma00 + &gamma10*sin_Func + &gamma20*cos_Func
            + &gamma30*Xw_it
            + &gamma01*TICi + &gamma02*Xb_i
            + xi_0i + epsilon_it;
 output;
```

```
        end;  * <-- end inner loop, go to next level-1 unit;
    end;  * <-- end outer loop, go to next level-2 unit;
run;
```

c) Study 2 script. The SAS script is largely the same as that in a) and in b), except that the cyclical components in the outcome is defined by two additional variables to account for the phase difference between cycles in the two series:

```
* cyclical components of the outcome;
sin_Func_DV=sin((2*CONSTANT('PI')*(time-PhaseD))/7);
cos_Func_DV=cos((2*CONSTANT('PI')*(time-PhaseD))/7);
```

So the Y series is generated as:

```
Yit=&gamma00
        + &gamma10*sin_Func_DV + &gamma20*cos_Func_DV + &gamma30*Xw_it
        + &gamma01*TICi + &gamma02*Xb_i
        + xi_0i + epsilon_it;
```

d) Study 3 script. The SAS script is the same as that in a) and in b), except that when cycles only exist in the outcome series, the parameters of the cyclical components (gamma_X10 and gamma_X20) in generating the time-varying predictor are set to 0, whereas when cycles only exist in the time-varying predictor series, the parameters of the cyclical components (gamma_10 and gamma_20) in generating the outcome variable are set to 0.