MULTIVARIATE BEHAVIORAL RESEARCH
2025, VOL. 60, NO. 3, 423–443

Evaluating Contextual Models for Intensive Longitudinal Data in the Presence of Noise

Anja F. Ernst[a], Eva Ceulemans[b], Laura F. Bringmann[a], and Janne Adolf[b]

[a]Department Psychometrics and Statistics, University of Groningen, Groningen, The Netherlands; [b]Department Quantitative Psychology and Individual Differences, KU Leuven, Leuven, Belgium

ABSTRACT

Nowadays research into affect frequently employs intensive longitudinal data to assess fluctuations in daily emotional experiences. The resulting data are often analyzed with moderated autoregressive models to capture the influences of contextual events on the emotion dynamics. The presence of noise (e.g., measurement error) in the measures of the contextual events, however, is commonly ignored in these models. Disregarding noise in these covariates when it is present may result in biased parameter estimates and wrong conclusions drawn about the underlying emotion dynamics. In a simulation study we evaluate the estimation accuracy, assessed in terms of bias and variance, of different moderated autoregressive models in the presence of noise in the covariate. We show that estimation accuracy decreases when the amount of noise in the covariate increases. We also show that this bias is magnified by a larger effect of the covariate, a slower switching frequency of the covariate, a discrete rather than a continuous covariate, and constant rather than occasional noise in the covariate. We also show that the bias that results from a noisy covariate does not decrease when the number of observations increases. We end with a few recommendations for applying moderated autoregressive models based on our simulation.

KEYWORDS

Moderated autoregression; fixed moderated time-series analysis; autoregressive models; measurement error

Introduction

Psychological research increasingly studies the dynamics with which emotions fluctuate over time within individuals (e.g., Hamaker et al., 2016; Kuppens et al., 2010; Kuppens & Verduyn, 2017). Inter-individual differences in these dynamics have been linked to psychological well-being and psychopathology (Brose et al., 2015a; Houben et al., 2015; Kuppens & Verduyn, 2017; van Roekel et al., 2018). Autoregressive (AR) models are currently widely used to quantify the dynamic properties of emotional states over time (Kuppens & Verduyn, 2017). In an AR model an observation at a given time-point is regressed on the observation at a previous time-point (Hamilton, 1994). A reason for the popularity of this model is that the corresponding regression parameter (AR coefficient) quantifies the propensity of affective states to resist change and persist over time, a concept that is coined emotional inertia (Houben et al., 2015; Kuppens & Verduyn, 2017). Further, an intercept is often included in the AR model to account for non-zero expected values.[1] Lastly, the innovation variance, the variance of the process residuals, is taken to reflect variability in the emotional process due to the influence of all factors that are not directly measured in the AR model (Hamaker et al., 2018).

Usually the AR model parameters (i.e., AR coefficient, intercept, and innovation variance) are assumed to be constant over time. However, emotions and their dynamics are not constant but likely change across different contexts. For instance, emotions dynamics might change over time (Bringmann et al., 2018; Lancee et al., 2022), across interpersonal contexts (Sels et al., 2022), during stressful or unpleasant situations (e.g., as induced in an experiment (Sels et al., 2020)), across different times of the day (Ernst et al., 2020), or might be altered suddenly by traumatic events (Simons et al., 2021). Recently, researchers are calling for such contexts to be sufficiently taken into account when modeling emotion dynamics (Dejonckheere et al., 2020; Lapate & Heller, 2020; Mestdagh & Dejonckheere, 2021). In an AR model, changes in emotion dynamics can be modeled through, for example, change-point detection methods (Albers & Bringmann, 2020; Cabrieto et al., 2018; Sels et al., 2022), regime-switching models (Chow et al., 2013; Crayen et al., 2017; Fuchs et al., 2017; Griffin & Li, 2016; Stifter & Rovine, 2015), threshold autoregressive models (Haan-Rietdijk et al., 2016), regression splines (Bringmann et al., 2018), or moderated AR models (MAR models henceforth, see Adolf et al., 2017; Bringmann et al., 2024; Ernst et al., 2020; Haslbeck et al., 2021; McNeish & Hamaker, 2020).

Here we focus on MAR models where observed covariates moderate the model parameters as this is one of the relatively simple and computationally light ways of accounting for changing dynamics (Adolf et al., 2017). These MAR models thus use a covariate, for instance stress at work, to predict changes in the emotion dynamics of a person within the AR framework. In this paper we focus on covariates that predict changes in the intercept and AR coefficients, as these are often the parameters of interests (Kuppens et al., 2010). Other parameters in the model could also be influenced by a covariate, for instance the innovation variance (e.g., Adolf et al., 2017; McNeish & Hamaker, 2020).[2]

A problem with tying changes in emotion dynamics to an observed covariate, such as work stress, is that changes in emotion dynamics are often caused by an unobserved covariate and we can observe only an imperfect measure of it that consequently contains measurement error. Measurement error can be caused by: (1) inaccurate recording of responses (e.g., entering an unintended response), or by (2) random error, as human responses to questions like "How demanding is your job at this moment?" are known to randomly fluctuate around the true value at that time. "The amount of measurement error variance in some measures used in psychological research is large, often in the neighborhood of 50% of the total variance of the measure" (Schmidt & Hunter, 1996, p. 200). In this paper we refer to covariates that contain measurement error as *noisy*. Noise can also arise when there is an omitted variable, that is when another unobserved covariate that is not accounted for also cause changes, and this unobserved covariate covaries with the included covariate. In this paper we focus exclusively on the measurement error case, but we will discuss the extension to the omitted variable case in the "Discussion" section.

For cross-sectional moderation models, various factors that influence their estimation accuracy have been studied, for discrete (Aguinis & Stone-Romero, 1997) and continuous covariates (Stone-Romero & Anderson, 1994). Of all these factors, noise in the covariates stands out as one of the most crucial, as it causes considerable bias in correlation and regression estimates, potentially leading researchers to erroneous conclusions (Liu & Salvendy, 2009). As a result, the power to detect moderation effects decreases rapidly as the amount of noise in the covariate and/or predictor variable increases (Aguinis, 1995; Dunlap & Kemery, 1988; Stone-Romero & Anderson, 1994). This decrease can already be substantial for very small amounts of noise (Aguinis, 1995; Dunlap & Kemery, 1988). Additionally, the power to detect moderation effects decreases as the sample size decreases (Aguinis & Stone-Romero, 1997), the range of the predictor variable is restricted (Aguinis, 1995; Aguinis & Stone-Romero, 1997), or there are unequal sample sizes across covariate based subgroups (Aguinis, 1995; Aguinis & Stone-Romero, 1997). Also, the artificial dichotomization of a continuous covariate can lead to a decrease in power (Stone-Romero & Anderson, 1994) and to an increase in Type I errors (Kang & Waller, 2005).

Crucially, these factors that lower the estimation accuracy of moderation effects interact with one another (i.e., they have non-additive effects) (Aguinis & Stone-Romero, 1997; Kang & Waller, 2005), often leading to detrimentally low estimation accuracy in empirically realistic settings. In light of these findings, many empirical studies have been shown to have had inadequate power to detect a moderation effect (Aguinis & Stone-Romero, 1997), illustrating the need to determine the estimation accuracy of moderation models under specific research conditions (i.e., for possible combinations of different factors that influence their estimation accuracy) (Dunlap & Kemery, 1988). This highlights the need for simulation studies to establish how well MAR models can be estimated in certain conditions when factors that influence their estimation accuracy, like noise, co-occur and potentially interact with factors that are specific to time-series data, such as a covariate being measured repeatedly over time.

For intensive longitudinal data it has been established by previous research that noise in the outcome variable (which constitutes a predictor) of an AR

[2]McNeish and Hamaker (2020) address between-individual differences in the innovation covariance matrix through a multilevel model, in this paper, however, we focus on within-individual differences.

model causes parameter estimates to be biased (Schuurman et al., 2015). Schuurman et al. (2015) found that estimates of the AR parameter already exhibit bias for relatively small proportions of noise in the outcome variable (i.e., 13%). The effects of a noisy *covariate* in MAR have, however, not been extensively studied thus far. This is concerning because recent research suggests that applied researchers often do not consider the psychometric properties of their intensive longitudinal data (Vogelsmeier et al., 2024) and also because the research on cross-sectional moderation models suggests that the estimation accuracy of MAR models will be severely impacted by noise in the covariate (Aguinis, 1995; Dunlap & Kemery, 1988; Stone-Romero & Anderson, 1994).

The aim of this paper is to investigate to what extent the estimation accuracy, assessed in terms of bias and variance, of different MAR models will be impacted when the covariate is a noisy predictor of changes in the dynamics. In addition to the amount of noise, we also investigate how other covariate characteristics influence the estimation accuracy of MAR models when they co-occur with noise. We investigate characteristics such as the format of the covariate (i.e., discrete or continuous), the switching frequency of the covariate (e.g., if the covariate measurements that are closely spaced in time are similar to each other), and the time-structure of the noise in the covariate (i.e., if noise is present constantly, at all measurements or only occasionally, at some measurements). This will add to the recent line of research on how well (M)AR-type models can be estimated under realistic conditions (e.g., Adolf et al., 2017; Ariens et al., 2023).

Overview

This paper is organized as follows, first we illustrate a standard AR model, followed by extensions to MAR models and the underlying assumptions of these models. Second, we illustrate the estimation of MAR models in the case of a noisy covariate. Third, we describe the covariate characteristics that we will investigate in this paper. Fourth, we present our simulation study where we examine the estimation accuracy of different MAR models across many empirically relevant situations. Fifth, we present our simulation results on how estimation accuracy is influenced by the different covariate characteristics that we have listed above. We end with an evaluation of the MAR models we have illustrated by discussing the implications of their underlying statistical assumptions, and formulating recommendations for applying them based on our simulation study.

Autoregressive models

In the following we describe the different (M)AR models we will consider in this paper. Throughout this paper we will exclusively consider (M)AR models with a time lag of one, often denoted (M)AR(1), henceforth we drop the notation that indicates the time lag for simplicity. AR models describe the dynamics of an emotion of interest, for example a person's mood denoted by η_t, over time-points t with $t = 1, \ldots, T$. We also show the extensions to MAR models where we include the covariate X to account for contextual influences by letting the intercept and/or the autoregression be moderated by covariate X. X could be any time-varying variable, for instance, the person's work stress as rated on a scale from one to 100, or an experimental condition that is experienced as either pleasant or stressful (dummy coded). The effect of the covariate can be included in a contemporaneous way, with x_t effecting outcomes at time-point t, or in a lagged way, with x_{t-1} effecting outcomes at time-point t. Here we stick to a lagged effect, we thus use work stress at a previous time-point, x_{t-1}, as a covariate when predicting a person's current mood, η_t.

Basic AR model

In a basic AR model the emotion dynamics of a person are modeled by predicting a person's mood on a given measurement, η_t, by their mood at the previous measurement η_{t-1}. This model implies that values that are closely spaced in time will be more similar to each other than values that are further apart. The basic AR model can be written as

$$\eta_t = \alpha + \rho\eta_{t-1} + \zeta_t \quad \zeta_t \sim N(0, \sigma^2)$$

where α represents the intercept, and ρ represents the AR coefficient which indicates the direct influence of η_{t-1} on η_t (i.e., the carry-over effect or inertia). ζ_t represents the innovations which indicate the process residuals. These innovations are passed to future time points through ρ. All (M)AR models make the following assumptions. Innovations ζ_t are assumed normally distributed with mean zero and variance σ^2. Further, innovations are assumed serially independent and independent of all predictors that appear with them in the regression equation (i.e., the previous value of the outcome variable, η_{t-1}) (Wooldridge, 2009, p. 351). Additionally, the underlying process is assumed weakly stationary. For a Gaussian process this implies that the mean (which is given by $\mu_\eta = \frac{\alpha}{(1-\rho)}$), variances, and autocovariances[3] of the outcome variable are assumed stable over time (Hamilton, 1994). This assumption implies a restriction on the AR parameter to be no greater than $|1|$ (Hamilton, 1994). Further, the measurements are assumed equidistant with equal time-intervals between them.

Including contextual events as a moderator for the intercept and autoregression

A MAR model where the value of the intercept and the autoregression are influenced by a covariate can be specified as follows. We denote the following model as the IntAR-MAR model

$$\text{Structural model}: \eta_t = \alpha^t + \rho^t \eta_{t-1} + \zeta_t \quad \zeta_t \sim N(0, \sigma^2) \tag{2}$$

$$\text{Moderation}: \quad \rho^t = \rho + \beta_\rho x_{t-1} \tag{3}$$

$$\alpha^t = \alpha + \beta_\alpha x_{t-1}, \tag{4}$$

where α^t represents the intercept at time-point t, α represents the value of the intercept when $x_{t-1} = 0$, and β_α represents the effect of the covariate at the previous time-point, x_{t-1}, on the intercept at the current time-point t. ρ^t is the AR coefficient at time t, ρ represents the value of the AR coefficient when $x_{t-1} = 0$, and β_ρ represents the effect of the covariate at the previous time-point, x_{t-1}, on the AR coefficient at the current time-point t. MAR models assume conditional stationarity, thus they make the same assumptions as AR models, except all assumptions are conditional on the covariate x_{t-1} (Adolf et al., 2017).

Including contextual events as an influence for the intercept

Besides the full IntAR-MAR model, a special case of the model could be specified where the covariate influences only the intercept and consequently, β_ρ is equal to zero. In the following we denote such a model the Int-MAR model (usually referred to as ARX(1) model, see e.g., Lütkepohl, 2005, Chapter 10)

$$\text{Structural model}: \eta_t = \alpha^t + \rho \eta_{t-1} + \zeta_t \quad \zeta_t \sim N(0, \sigma^2) \tag{5}$$

$$\text{Moderation}: \alpha^t = \alpha + \beta_\alpha x_{t-1}, \tag{6}$$

Including contextual events as a moderator for the autoregression

Also, a special case of the IntAR-MAR model could be specified where the covariate influences only the autoregression and consequently β_α is equal to zero. In the following we denote such a model as the AR-MAR model

$$\text{Structural model}: \eta_t = \alpha + \rho^t \eta_{t-1} + \zeta_t \quad \zeta_t \sim N(0, \sigma^2) \tag{7}$$

$$\text{Moderation}: \quad \rho^t = \rho + \beta_\rho x_{t-1}, \tag{8}$$

Estimation

Despite accommodating time-series data, the MAR models shown above are simply moderated regression models (Cohen et al., 2003). These models can thus be estimated through ordinary least squares (Lütkepohl, 2005), maximum likelihood estimation (Adolf et al., 2017; Lütkepohl, 2005), or Bayesian estimation (Speyer et al., 2024). Software that can be used for the maximum likelihood estimation of MAR models include dynr (Ou et al., 2019) and OpenMx (Adolf, 2023; Neale et al., 2016) in R. For some MAR models, the vars R-package can be used for ordinary least squares estimation (i.e., for the Int-MAR model, see Pfaff, 2008). Also, some MAR models can readily be estimated in Mplus through Bayesian estimation (Koval & Kuppens, 2012; McNeish & Hamaker, 2020; Muthén & Muthén, 2013).[4] Alternatively, Stan can be used for Bayesian estimation of MAR models (Stan Development Team, 2024). In our simulation we employ maximum likelihood estimation. However, our findings of a noisy covariate causing bias generalize to all other estimation methods as well, because this bias is due to model misspecification. Hence, this bias is independent of the estimation method (Frost & Thompson, 2000), unless methods that are robust toward misspecification were used, as discussed in the Discussion section.

Factors that influence the estimation accuracy of a MAR model

The different MAR models that were introduced above are depicted as path diagrams in the left panels of Figure 1; here moderated parameters are shown in

[3]Autocovariances describe the covariance of the outcome variable and its own lagged value.

[4]Mplus employs the within-person mean specification as mentioned in the Discussion section and as shown in in Equations A.9–A.12 in Supplemental material, supplemental content D.

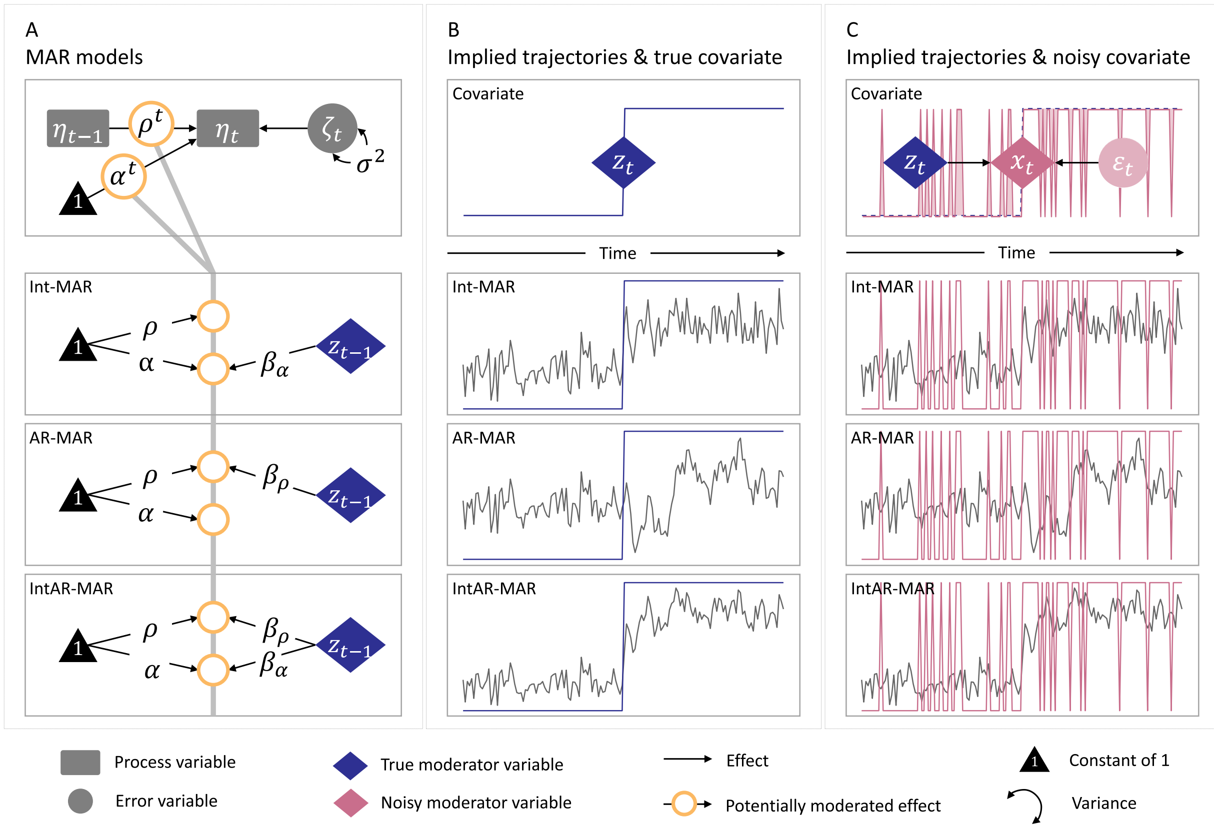

Figure 1. MAR models, their behavior and covariates. Panel A displays path diagrams of all MAR models considered and panels B and C show model-implied trajectories combined with the trajectories of the true and noisy observed covariate respectively. The figure is restricted to the case of a slow-switching covariate with occasional discrete noise.

circles (Curran & Bauer, 2007). An example of how a person's mood might change based on each of these different MAR models as true data generating model is visualized in the middle column of Figure 1. These panels also show the true covariate that influences the emotion dynamics of mood in this example: Changes in the true covariate visibly coincide with changes in the emotion dynamics. The right panels of Figure 1 show again the person's mood based on these different MAR models but now next to a noisy observed covariate. Using the observed covariate, it becomes harder to distinguish whether changes in the covariate coincided with changes in the emotion dynamics. In practice, this likely leads to decreased estimation accuracy of various model parameters. The aim of the following simulation is to investigate to what extent the estimation accuracy of different MAR models will be impacted by noise in the covariate. We also aim to investigate how other properties of the covariate or of the noise impact the decrease in estimation accuracy that is caused by noisy covariates. This will extend the recent line of research on how well (M)AR-type models can be estimated under realistic conditions (e.g., Adolf et al., 2017; Ariens et al., 2023).

In Figure 1, the true covariate persists in phases of similar observations for several time-points before 'switching' to a different phase of similar observations. Such switching behavior is common for covariates in empirical applications because, often contextual events persist over time so that contextual events that are closely spaced in time tend to be similar to each other. When recording the work-stress that a person experiences, for instance, one might observe that many high-stress observations occur one after another. Depending on the sampling frequency (i.e., the time gap between measurements) and the nature of the covariate, such phases of similar observations can persist for many time-points, resulting in a slow-switching covariate, or for few time-points, resulting in a fast-switching covariate. Fast-switching covariates show low temporal persistence while slow-switching covariates show high temporal persistence. The switching frequency of a covariate is another characteristic (next to the amount of noise) that can influence the estimation accuracy of MAR models. For example, Adolf et al. (2017) showed that in the absence of noise, a slower switching frequency of the covariate caused higher bias in MAR models.

Similarly, (Ariens et al., 2023) demonstrated that in the absence of noise, serial dependence (i.e., due to autoregression or trending) in the covariate can result in larger standard errors for MAR models. In our simulation we will investigated to what extent different switching frequencies will impact the estimation accuracy when they co-occur with noise in the covariate.

Instead of a discrete covariate like in Figure 1, we could also consider a covariate with a different format, like a continuous covariate. The format of the covariate might also influences the estimation accuracy of MAR models, especially when the covariate contains noise; because a noisy continuous covariate can contain continuous noise, while a noisy discrete covariate can contain only discrete noise. We, therefore, consider different formats of the covariate (i.e., discrete and continuous covariates) in our simulation.

In Figure 1, noise in the covariate occurs occasionally with only certain measurements of the covariate including noise and other measurements being noise-free. Alternatively, if a variable is measured in a continuous manner, a given percentage of noise might be present at all measurements of the covariate, leading thus to constant noise in the covariate. To find if the time-structure of the noise influences the estimation accuracy, we include different time-structures of the noise in our simulation.

Expectations

We expect the following main and interaction effects: (1) In the presence of noise, the influence of the covariate (i.e., $\hat{\beta}_\rho$ and/or $\hat{\beta}_\alpha$) will be biased toward zero and this effect will become more severe with larger amounts of noise. We expect this because, in the presence of noise, regression coefficients are always biased toward zero, provided that only a single predictor variable is used and that the noise is unbiased and independent of the true value of the outcome variable (Frost & Thompson, 2000). This effect is referred to as regression dilution bias (MacMahon et al., 1990). (2) The bias toward zero of $\hat{\beta}_\rho$ and $\hat{\beta}_\alpha$ will lead to a bias away from zero (thus leading to larger absolute values) of $\hat{\rho}$ and $\hat{\alpha}$. (3) the autoregression will be most severely over-estimated in the slow-switching covariate conditions. This expectation arises because in all our simulation conditions, the covariate persists in phases (e.g., the covariate equals 1 for a number of successive time-points and then switches to be equal to 0 for a number of successive time-points). This leads to a greater similarity between observations, denoted as η_t, that occur around similar times due to comparable covariate influences. In the presence of noise, it is expected that the impact of the covariate (expressed as $\hat{\beta}_\rho$ and/or $\hat{\beta}_\alpha$) will be underestimated. Consequently, this underestimation leaves the similarity between closely spaced η_t unexplained. The unexplained similarity, which is due to comparable covariate influences, is then erroneously interpreted as heightened autoregression in η_t (indicated by $\hat{\rho}$). This overestimation of autoregression is anticipated to be more pronounced in the case of slow-switching covariates compared to fast-switching covariates. This distinction arises from the fact that, under slow-switching covariates, closely spaced observations η_t are similar to one another for longer. Therefore, if the influence of the covariate is underestimated in these conditions, the resulting unexplained longer similarity between successive observations leads to a greater overestimation of the autoregression.

Model assumptions

As stated above, (MA)R models assume that the innovations, ζ_t, are assumed independent of any of the predictors that appear with them in the regression equation (i.e., contemporaneous exogeneity, see Wooldridge (2009, p. 351) or Lütkepohl (2005, p. 389)). In our case this implies that ζ_t should be independent of x_{t-1} and η_{t-1}. If instead of the lagged effect the contemporaneous effect of X, x_t, would be included, ζ_t would have to be independent of x_t. Such contemporaneous exogeneity is sufficient for the consistent estimation of model parameters (Pesaran, 2015; Wooldridge, 2009) though stricter versions and more theoretical definitions of exogeneity exist (see e.g., Hendry (1995, Chapter 5) or Engle et al. (1983)). The assumption of contemporaneous exogeneity will be violated, for instance, when the observed covariate contains measurement error and the measurement error is correlated with the observed covariate[5] (see e.g., Wooldridge (2002) pp. 71—76 or our illustration in the Supplemental material, supplemental content A). Alternatively, the assumption of contemporaneous exogeneity will be violated when a variable is omitted that influences the outcome and that is correlated with the observed covariate (see e.g., Wooldridge (2002), pp. 50—51 or our illustration in the Supplemental material, supplemental content A). We will investigate the effect of this assumption violation in our simulation study where this assumption will be

[5]The measurement error contained in an observed covariate will always be correlated with the observed covariate, provided that the measurement error is uncorrelated with the true covariate (Wooldridge, 2002, pp. 73–76).

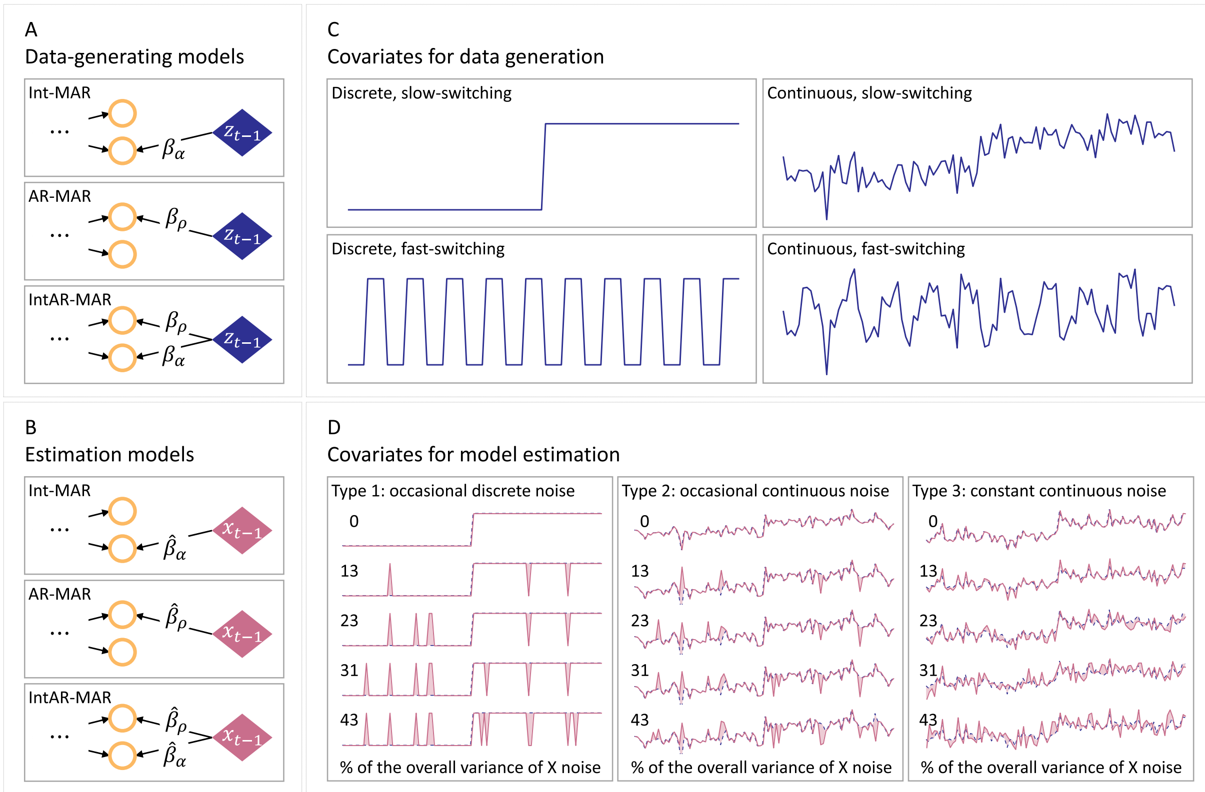

Figure 2. Simulation design aspects. Panels A and B display partial path diagrams of the MAR models used for data generation and estimation respectively. Panels C and D show the types of true and observed covariates used, in this case restricted to the case of 100 time points.

violated because of measurement error in the observed covariate.

Simulation

We conduct a simulation study to assess how the estimation accuracy of MAR models is impacted by noise in the covariate under various relevant conditions. In this simulation, we distinguish between the observed covariate X, which is used in the estimation of the model, and the true covariate Z, which is used in the data generation of the model. Additionally, the observed covariate X is generated to be a noisy measure of the true covariate Z.

We investigate to what extent estimation accuracy depends on seven factors: (1) the type and number of parameters that are influenced by the covariate (i.e., the intercept, the autoregression, or the intercept and the autoregression), (2) the number of time-points, and (3) the effect size of the influence of the covariate. Two factors reflect how the true covariate for the data-generating models are generated (i.e., the true covariate Z): (4) the switching frequency of the covariate (we consider slow-switching and fast-switching covariates) and, (5) the format of the covariate (i.e., discrete or continuous). Two factors reflect how the observed covariate for the estimation models are generated (i.e., the observed covariate X): (6) the time-structure of the noise (which depends on the format of the covariate), we consider covariates that contain occasional noise or constant noise, and (7) the amount of noise in the covariate. In the following paragraphs we describe these seven simulation factors in more detail. Figure 2 shows an overview of our simulation design.

Simulation method

Across all ($3 \times 3 \times 3 \times 2 \times 3 \times 5 = 810$ in total) simulation conditions, the number of replications per condition was set to 500, the innovation variance, σ^2, was set to 0.5, the baseline autoregression, ρ, equaled 0.2,[6] and the baseline intercept, α, equaled 0. The following simulation factors were varied in a completely crossed design.

[6]In our simulation this value of 0.2 pertains to the baseline autoregression, ρ, thus for many of the time-points in our simulation the value of the 'overall' autoregression, ρ^t, is higher than 0.2 and is given by Equation 3.

1. The type and number of parameters that are influenced by the covariate. This factor pertains to the model that was used to generate the data. Data was generated according to one of the three moderation models specified in the Introduction section (i.e, either Int-MAR, AR-MAR, or IntAR-MAR). Because we always fitted the 'true' model to the data, the data generation model is always identical to the model that was then subsequently fitted to the data:
   - Data was generated according to the Int-MAR model (Equation (5)), this model was then also fitted to the data. Estimation accuracy was evaluated in these conditions for ρ, β_α, and α.
   - Data was generated according to the AR-MAR model (Equations (7) and (8)),[7] this model was then also fitted to the data. Estimation accuracy was evaluated in these conditions for ρ, β_ρ, and α.
   - Data was generated according to the IntAR-MAR model (Equations (2) and (3)).[8] This model was then also fitted to the data. Estimation accuracy was evaluated in these conditions for ρ, β_ρ, β_α, and α.
2. The total number of observations: 100, 200, or 500.
3. The effect sizes of the covariate influence, expressed in values of β_α (if specified in the data generation model) and/or β_ρ (if specified in the data generation model):
   - Low effect sizes of covariate influence, with $\beta_\rho = 0.15$ and/or $\beta_\alpha = 0.36$;
   - Medium effect sizes of covariate influence, with $\beta_\rho = 0.20$ and/or $\beta_\alpha = 0.72$;
   - High effect sizes of covariate influence, with $\beta_\rho = 0.25$ and/or $\beta_\alpha = 1.08$.

These effect sizes of β_ρ and β_α are not equivalent to one another. Rather we chose the effect sizes for β_ρ in such a way that they would result in a broad range of values for ρ^t (see Equation (3)) that correspond to values that are observed in practice for emotion dynamics and in values that do not exceed $|1|$. For example, in conditions with $\beta_\rho = 0.25$, ρ^t is equal to 0.2 when $x_{t-1} = 0$ and to 0.45 when $x_{t-1} = 1$. The effect sizes for β_α were chosen so that they correspond respectively to an increase in the intercept, α, of 0.5, 1, or 1.5 standard deviations of the time-series when the covariate is equal to 0.[9] We express the size of the effects for β_α in measures of the standard deviations of the time-series because it takes the autoregression and innovation variance into account. This facilitates comparison of the effect sizes employed in our simulation with the effect sizes observed in data with different values for autoregression and innovation variance.

4. The switching frequency of the covariate:
   - *Fast-switching covariate* where the covariate was switching between 'condition A' and 'condition B' (see details below) every five time-points;
   - *Slow-switching covariate* where the covariate was switching between 'condition A' and 'condition B' every 50 time-points.

   We thus consider scenarios where true covariate Z describes contextual events that persist for five observations (fast-switching covariate) or for 50 observations (slow-switching covariate).

5.&6. The type of covariate. This combines two of our simulation factors because these factors depend on each other. Combining (5) the format of the covariate (i.e., either a discrete or a continuous covariate) and (6) the time-structures of the noise, results in the following three types of covariate:
   - A discrete covariate with noise at some random time-points (*Type 1: Occasional discrete noise covariate*);
   - A continuous covariate with noise at some random time-points (*Type 2: Occasional continuous noise covariate*);
   - A continuous covariate with noise at all time-points (*Type 3: Constant continuous noise covariate*).

7. The amount of noise in the covariate. To evaluate how the amount of noise in the covariate influences the estimation accuracy of MAR models, we add noise (i.e., measurement error) to the covariate in the following way. We first generate a true covariate Z, which is used to generate data. We then add noise to Z in order to create a noisy covariate X that is then used in the model estimation as the observed covariate. Adding noise in this fashion simulates the scenario where the covariate is measured inaccurately and consequently contains measurement error. As can be

[7]When a randomly generated value for the covariate was so high/low as to imply a locally non-stationary process (i.e., implying a value for ρ^t that is larger than $|1|$), the covariate was generated again.
[8]See footnote 7.

[9]The standard deviation of a time-series with an AR coefficient of 0.2 and an innovation variance of 0.5 equals $\sqrt{\frac{0.5}{1-0.2^2}} \approx 0.72$.

seen when comparing the path diagrams for the data generating models and the estimation models that are shown in Figure 2, their difference lies in the estimation models using the noisy observed covariate X in place of the true covariate Z. The amount of noise contained in observed covariate X was manipulated on five levels in terms of the proportion of noise ($\frac{var(Noise)}{var(Z+Noise)} = \frac{var(Noise)}{var(X)}$):

- No noise in the observed covariate X (i.e., observed covariate X is identical to true covariate Z);
- 13% of the observed covariate X is noise;
- 23% of the observed covariate X is noise;
- 31% of the observed covariate X is noise;
- 43% of the observed covariate X is noise;

These proportions of noise were chosen because they correspond to simulation conditions in Schuurman15 who investigated the effect of including measurement error in an AR model that does not include a covariate. In order to ensure that all three different types of covariates employed in our simulation will have these proportions of noise, the 15 conditions that result from crossing simulation factors (5 & 6) 'The type of covariate' with (7) 'The amount of noise', are simulated as follows:

- Occasional discrete noise covariate (Type 1): Z is equal to zero for all observations from 'condition A', Z is equal to one for all observations from 'condition B'. The amount of noise that is added to Z to create X is manipulated on five levels:
  - No noise is added;
  - 3% of random values in Z are changed (i.e., from 0 to 1 or from 1 to 0);
  - 6% of random values in Z are changed;
  - 8% of random values in Z are changed;
  - 11% of random values in Z are changed;[10]
- Occasional continuous noise covariate (Type 2): The value of Z for any observation from 'condition A' is drawn from $N(0, 0.1)$,[11] the value of Z for any observation from 'condition B' is drawn from $N(1, 0.1)$.[12] The amount of noise that is added to Z to create X is manipulated on five levels:
  - No noise is added;
  - 4% of randomly selected values in Z are changed (i.e., when the selected value is in 'condition A', it is replaced with a value generated from $N(1, 0.1)$; when the selected value is in 'condition B', it is replaced with a value generated from $N(0, 0.1)$);
  - 7% of randomly selected values in Z are changed;
  - 9% of random values in Z are changed;
  - 13% of random values in Z are changed;[13]
- Constant continuous noise covariate (Type 3): the value of Z for any observation from 'condition A' is drawn from $N(0, 0.1)$, the value of Z for any observation from 'condition B' is drawn from $N(1, 0.1)$. The amount of noise that is added to Z to create X is manipulated on five levels:
  - No noise is added;[14]
  - To each value of Z, a value that is drawn from $N(0, 0.052)$ is added (this corresponds to $\sim$13% of the total variance of X, details can be found in the Supplemental material, supplemental content B);
  - To each value of Z, a value that is drawn from $N(0, 0.105)$ is added (this corresponds to $\sim$23% of the total variance of X);
  - To each value of Z, a value that is drawn from $N(0, 0.158)$ is added (this corresponds to $\sim 31\%$ of the total variance of X);
  - To each value of Z, a value that is drawn from $N(0, 0.265)$ is added (this corresponds to $\sim 43\%$ of the total variance of X).

Figure 2 displays the covariates that result from our simulation conditions. The top of Figure 2 shows the different measurement formats and switching frequencies of the true covariate Z, the bottom parts shows the observed covariate X that is created by adding noise to true covariate Z, resulting in Occasional discrete noise covariates, Occasional continuous noise covariates, and Constant continuous noise covariates with varying proportions of noise. Data was generated with the fmTSA package (Adolf, 2023) in R. Data was analyzed as a linear discrete-time model with the dynr package (Ou et al., 2019), which employs maximum likelihood estimation. Here a Kalman Filter is used to construct the log-likelihood function, known as the prediction error decomposition function (Ou et al., 2019). Optimization of this likelihood function yields maximum likelihood estimators for the parameters of interest. Estimation with dynr requires starting values for all estimators, in line with recommendations by Liu et al. (2021) we picked starting values of 0.1. All

[10]We selected these percentages because they result in average noise proportions of roughly 13%, 23%, 31%, or 43%, respectively.
[11]A normal distribution with mean 0 and a *variance* of 0.1.
[12]A normal distribution with mean 1 and a *variance* of 0.1.
[13]We selected these percentages because they result in average noise proportions of roughly 13%, 23%, 31%, or 43%, respectively.
[14]The no noise conditions for Occasional continuous noise covariates and Constant continuous noise covariates are thus identical.

the R scripts used to generate data, carry out the simulations, and analyze the data are provided on the project's OSF page.[15]

Performance measures

We evaluate estimation accuracy in terms of bias, assessed as the mean difference between parameters and estimators across all simulation replications, and variance, assessed as the variance of estimators obtained across all simulation replications (see Hastie et al., 2009, p. 24). To be able to better evaluate the size of the bias, we employ the relative bias. For this we convert the bias to be expressed in measures of the true underlying effect size: $\frac{1}{\#\text{Replications}}\sum_{r=1}^{\#\text{Replications}}\frac{\hat{\theta}_r-\theta_r}{\theta_r}$ (Hoogland & Boomsma, 1998). Further, to quantify the uncertainty in our simulation results due to random sampling of replications, we bootstrap the sampling distribution of bias estimates with 500 bootstrap replicates. We calculate the standard error of the sampling distribution of the bias for all estimators in all 810 cells of our completely crossed design. Subsequently, we calculate the average bootstrapped standard error for all main effects by averaging all the bootstrapped standard errors of the cells that are contained in that condition (c): $\frac{1}{\#\text{Cells}}\sqrt{\sum_{c=1}^{\#\text{Cells}}\text{Standard Error}_c^2}$.

Model misspecification

We assess the performance of the different MAR models in case of (almost) correct model specification, meaning that the data was generated according to one of the MAR models specified in the introduction (e.g., the Int-MAR model) and were then subsequently fitted to this same model (i.e., the Int-MAR model). Despite this, in the conditions in which there is noise in the covariate, there will be slight model misspecification because the covariate being used to fit the model (i.e., X) is not the true covariate (i.e., Z).

Simulation results

Here we present the results for each of the three MAR models. In the results we exclude estimates of models that did not converge. The non-convergence rate was 0.21%, 0.14%, and 1.00% respectively for the Int-MAR, AR-MAR, and IntAR-MAR models. Additionally, we excluded three models (out of 405000) because the absolute value of their estimate for the baseline autoregression exceeded 1.[16]

Figure 3 shows the bias for each type of covariate condition crossed with the amount of noise. Figure 4 shows the bias for the different switching frequencies crossed with the amount of noise. These figures thus show the interaction effect of type of covariate with amount of noise and the interaction effect between switching frequency and the amount of noise. Tables 1–3 display the bias and variance across all simulation factors per MAR model, thus showing the main effects for each simulation factor. It can be clearly seen in these figures and tables, that bias increases rapidly with the amount of noise. When noise is absent, the bias is small for estimators of all parameters across all MAR models, which can be seen by the means falling on the gray line for these conditions in Figures 3 and 4. For the AR estimator $\hat{\rho}$ for instance, the average bias over all conditions with zero noise is equal to $-.011$ across all three models (as can be seen in Tables 1 through 3). In terms of relative bias, this corresponds to 6% of the true effect size, which is slightly above the cutoff of 5% that Hoogland and Boomsma (1998) use to indicate a small relative bias. When noise increases, however, bias increases rapidly for all estimators across all MAR models leading to substantial bias for all estimators in high-noise conditions. For the AR estimator $\hat{\rho}$ for instance, the average bias over all conditions with a noise proportion of 43% is equal to 0.066, 0.022, or 0.120 for the Int-MAR, AR-MAR, and IntAR-MAR models respectively (this can be seen in Tables 1–3). When converting these average biases to relative biases (Hoogland & Boomsma, 1998), they correspond to 33%, 11%, or 60% of the true effect size of the AR estimator, respectively. An exception to this pattern of increasing bias with increasing noise are the estimators for the intercept ($\hat{\alpha}$) in the AR-MAR model, in the AR-MAR model, these estimators display small bias even in high noise conditions, as can be seen in the plot in the first row, second column of Figure 3. Tables 4–6 show the average bootstrapped standard errors for the bias estimates, which quantify the uncertainty due to random sampling of replications in our simulation results for bias. These standard errors are relatively small (all are below 0.001), we derive that our 500 replications provide sufficiently accurate estimates of bias.

In line with our expectations, the following general patterns become visible across the figures and tables for all three MAR models, in the presence of noise:

[15] https://osf.io/sjtfk/?view_only=bd7dcc62fd214effa2b2e61e588a113d

[16] These three excluded models were all IntAR-MAR models in conditions in which the number of observations was equal to 500.

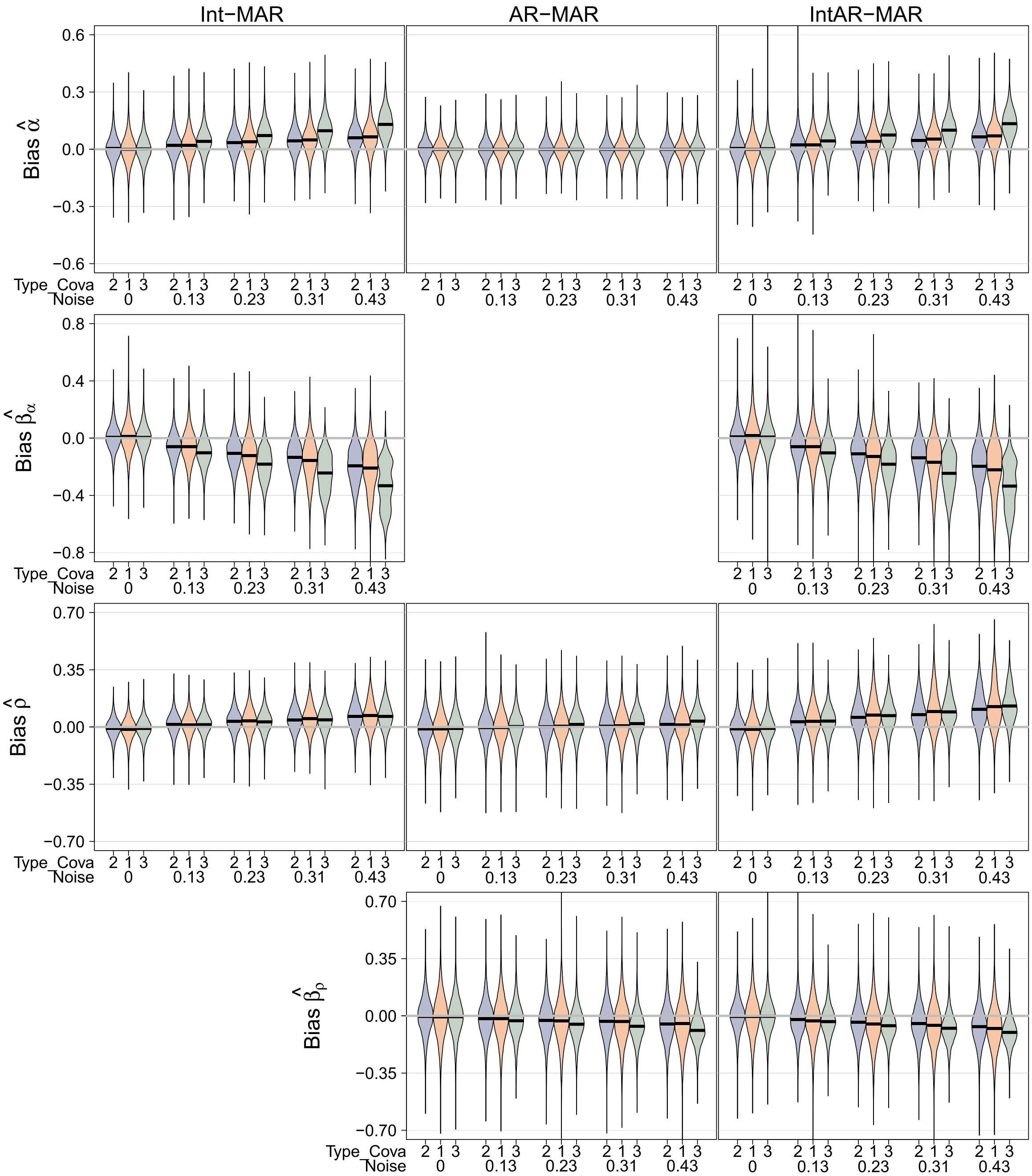

Figure 3. Interaction effects for type of covariate crossed with the amount of noise.
Note. The figure shows distances between parameter and estimators (bias). Type 1: Occasional discrete noise, Type 2: Occasional continuous noise, Type 3: Constant continuous noise. The first column shows the results of the Int-MAR model, the second column the results of the AR-MAR model, and the third column the results of the IntAR-MAR model. A black box indicates the area within one standard error from the mean, a gray line highlights zero.

The baseline intercept estimator ($\hat{\alpha}$) and the baseline AR estimator ($\hat{\rho}$) exhibit positive bias, the covariate influence estimators ($\hat{\beta}_\alpha$ and $\hat{\beta}_\rho$) exhibit negative bias. These general patterns are visible in Figures 3 and 4 by looking at a row of plots, which show the results of the same estimators across the different MAR models, and by observing that the black lines, which indicate the mean of estimators across different conditions, fall consistently below or consistently above the gray zero line.

Figure 3 shows that in the presence of noise, there are differences in bias for the different covariate types: bias is lowest for the Occasional continuous noise covariate, followed by the Occasional discrete noise covariate, bias is highest for the Constant continuous noise covariate. This pattern is visible across all MAR

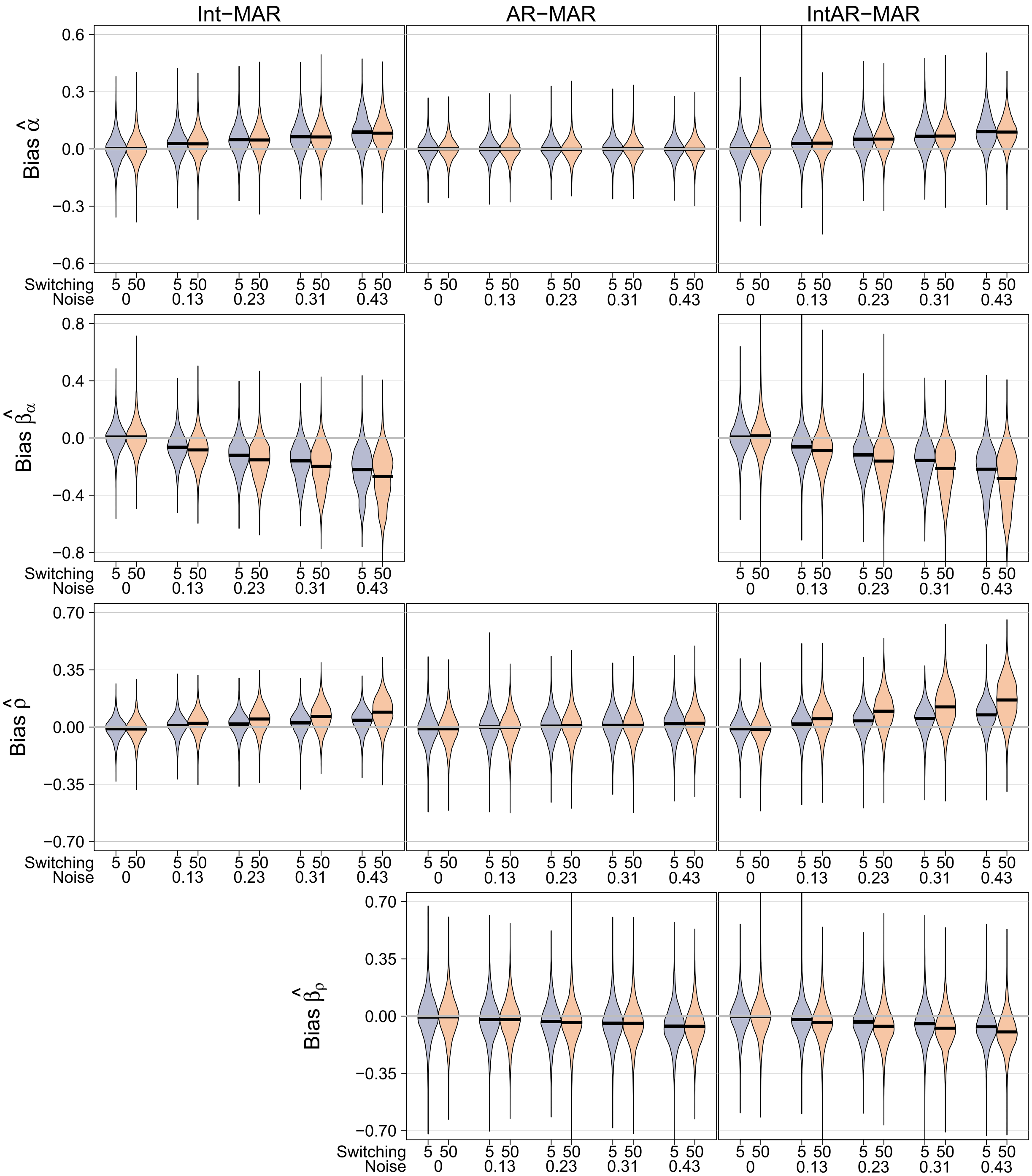

Figure 4. Interaction effects for switching frequency of the covariate crossed with the amount of noise.
Note. The figure shows distances between parameter and estimators (bias). The first column shows the results of the Int-MAR model, the second column the results of the AR-MAR, and the third column the results of the IntAR-MAR. A black box indicates the area within one standard error from the mean, a gray line highlights zero.

models in Figure 3. Regarding the different switching frequencies, the following differences between fast-switching and slow-switching covariates are visible in Figure 4: for the covariate influence estimators ($\hat{\beta}_\alpha$ and $\hat{\beta}_\rho$) and the AR estimator ($\hat{\rho}$), the slow-switching covariate has more bias than the fast-switching covariate. For the intercept estimator ($\hat{\alpha}$), there are no clear differences between the slow-switching and the fast-switching covariate in terms of bias. Because MAR models rely on several parameters, having less bias in one estimator does not equal less bias overall. Figure 4 shows that differences between fast-switching and slow-switching covariates are largest for the AR estimator ($\hat{\rho}$). In line with our expectations, the slow-switching

Table 1. Main effects for bias and variance across the simulation factors in the Int-MAR model.

	$\hat{\alpha}$ Bias	$\hat{\alpha}$ Var	$\hat{\beta}_\alpha$ Bias	$\hat{\beta}_\alpha$ Var	$\hat{\rho}$ Bias	$\hat{\rho}$ Var
100 Observations	.046	.011	−0.122	.084	.023	.010
200 Observations	.045	.006	−0.126	.074	.031	.006
500 Observations	.044	.003	−0.128	.068	.036	.003
Low ES	.027	.006	−0.053	.011	$\lvert x\rvert < .001$	.005
Medium ES	.048	.007	−0.121	.018	.029	.005
High ES	.061	.008	−0.203	.034	.061	.007
Type 2	.032	.006	−0.098	.076	.029	.006
Type 1	.035	.007	−0.107	.076	.032	.007
Type 3	.068	.008	−0.171	.071	.029	.006
Fast-switching	.046	.007	−0.112	.078	.017	.005
Slow-switching	.044	.007	−0.139	.072	.043	.008
0	.002	.006	.007	.099	−0.011	.005
0.13	.027	.006	−0.074	.077	.015	.005
0.23	.048	.006	−0.137	.063	.034	.006
0.31	.063	.006	−0.179	.055	.046	.006
0.43	.085	.007	−0.245	.046	.066	.008

Note. Bias indicates the mean difference between parameter and estimators, variance indicates the variance of estimators. Type 1: Occasional discrete noise, Type 2: Occasional continuous noise, Type 3: Constant continuous noise.

Table 2. Main effects for bias and variance across the simulation factors in the AR-MAR model.

	$\hat{\alpha}$ Bias	$\hat{\alpha}$ Var	$\hat{\rho}$ Bias	$\hat{\rho}$ Var	$\hat{\beta}_\rho$ Bias	$\hat{\beta}_\rho$ Var
100 Observations	−0.001	.006	−0.002	.017	−0.037	.030
200 Observations	$\lvert x\rvert < .001$	.003	.007	.008	−0.034	.015
500 Observations	$\lvert x\rvert < .001$	.001	.012	.003	−0.030	.007
Low ES	$\lvert x\rvert < .001$	.003	.002	.010	−0.026	.016
Medium ES	$\lvert x\rvert < .001$	.003	.005	.010	−0.033	.016
High ES	$\lvert x\rvert < .001$	.003	.010	.010	−0.042	.016
Type 2	$\lvert x\rvert < .001$	.003	.002	.009	−0.026	.017
Type 1	$\lvert x\rvert < .001$	.003	.002	.011	−0.027	.022
Type 3	−0.001	.003	.013	.009	−0.048	.014
Fast-switching	$\lvert x\rvert < .001$	.003	.005	.010	−0.033	.018
Slow-switching	$\lvert x\rvert < .001$	.003	.006	.010	−0.034	.017
0	$\lvert x\rvert < .001$	.003	−0.011	.010	−0.004	.019
0.13	$\lvert x\rvert < .001$	.003	−0.001	.010	−0.021	.017
0.23	−0.001	.003	.008	.010	−0.036	.017
0.31	$\lvert x\rvert < .001$	.003	.011	.009	−0.044	.017
0.43	−0.001	.003	.022	.009	−0.062	.016

Note. Bias indicates the mean difference between parameter and estimators, variance indicates the variance of estimators. Type 1: Occasional discrete noise, Type 2: Occasional continuous noise, Type 3: Constant continuous noise.

covariate has more bias than the fast-switching covariate in terms of the AR estimator ($\hat{\rho}$) and this estimator is over-estimated in the presence of noise.

In sum, Tables 1–3 show the following general main effects of bias that hold across all MAR models

Table 4. Averaged bootstrapped standard error for the bias (Int-MAR model).

	$\hat{\alpha}$	$\hat{\beta}_\alpha$	$\hat{\rho}$
100 Observations	0.0005	0.0006	0.0004
200 Observations	0.0003	0.0004	0.0003
500 Observations	0.0002	0.0003	0.0002
Low ES	0.0003	0.0004	0.0003
Medium ES	0.0003	0.0005	0.0003
High ES	0.0003	0.0005	0.0003
Type 2	0.0003	0.0005	0.0003
Type 1	0.0004	0.0005	0.0003
Type 3	0.0003	0.0004	0.0003
Fast-switching	0.0003	0.0004	0.0003
Slow-switching	0.0003	0.0004	0.0003
0	0.0005	0.0006	0.0004
0.13	0.0005	0.0006	0.0004
0.23	0.0004	0.0006	0.0004
0.31	0.0004	0.0006	0.0004
0.43	0.0004	0.0006	0.0004

Table 5. Averaged bootstrapped standard error for the bias (AR-MAR model).

	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\beta}_\rho$
100 Observations	0.0004	0.0006	0.0008
200 Observations	0.0002	0.0004	0.0006
500 Observations	0.0002	0.0003	0.0004
Low ES	0.0003	0.0005	0.0006
Medium ES	0.0003	0.0005	0.0006
High ES	0.0003	0.0005	0.0006
Type 2	0.0003	0.0005	0.0006
Type 1	0.0003	0.0005	0.0007
Type 3	0.0003	0.0004	0.0005
Fast-switching	0.0002	0.0004	0.0005
Slow-switching	0.0002	0.0004	0.0005
0	0.0003	0.0006	0.0008
0.13	0.0003	0.0006	0.0008
0.23	0.0003	0.0006	0.0008
0.31	0.0003	0.0006	0.0008
0.43	0.0003	0.0006	0.0007

Table 3. Main effects for bias and variance across the simulation factors in the IntAR-MAR model.

	$\hat{\alpha}$ Bias	$\hat{\alpha}$ Var	$\hat{\beta}_\alpha$ Bias	$\hat{\beta}_\alpha$ Var	$\hat{\rho}$ Bias	$\hat{\rho}$ Var	$\hat{\beta}_\rho$ Bias	$\hat{\beta}_\rho$ Var
100 Observations	.050	.011	−0.121	.093	.054	.020	−0.048	.026
200 Observations	.048	.006	−0.129	.078	.061	.012	−0.045	.013
500 Observations	.047	.004	−0.133	.071	.064	.008	−0.041	.007
Low ES	.029	.006	−0.052	.012	.015	.010	−0.028	.016
Medium ES	.052	.007	−0.122	.023	.059	.011	−0.043	.014
High ES	.064	.008	−0.211	.051	.107	.015	−0.064	.014
Type 2	.035	.006	−0.099	.080	.053	.012	−0.035	.014
Type 1	.038	.007	−0.112	.083	.063	.017	−0.043	.020
Type 3	.071	.008	−0.172	.076	.064	.011	−0.055	.012
Fast-switching	.048	.008	−0.109	.082	.035	.009	−0.034	.014
Slow-switching	.048	.007	−0.146	.079	.085	.017	−0.055	.016
0	.002	.006	.011	.110	−0.011	.008	−0.003	.017
0.13	.030	.006	−0.074	.081	.034	.009	−0.029	.015
0.23	.051	.006	−0.139	.066	.068	.011	−0.049	.014
0.31	.067	.006	−0.184	.057	.088	.013	−0.061	.014
0.43	.090	.007	−0.251	.048	.120	.015	−0.081	.014

Note. Bias indicates the mean difference between parameter and estimators, variance indicates the variance of estimators. Type 1: Occasional discrete noise, Type 2: Occasional continuous noise, Type 3: Constant continuous noise.

and almost all estimators: (1) In terms of bias there is no clear main effect for the number of observations. This makes sense as the bias observed in our simulation is not finite sample bias but instead is related to the noise (i.e., misspecification of the model), thus this bias will be present asymptotically. The variance of estimators, however, consistently decreases with higher numbers of observations. (2) Bias increases with the effect size of the covariate influence, the higher the covariate influence the higher the bias. (3) Bias is lowest for the Occasional continuous noise covariate, followed by the Occasional discrete noise covariate, bias is highest for the Constant continuous noise covariate. (4) Bias is higher for the slow-switching covariate than for the fast-switching covariate. (5) Bias is small in the conditions without noise but increases with noise, leading to very substantial bias across all estimators in the conditions with high noise. A small proportion of noise equal to 13% is already enough for a notable increase in bias for all estimators (with exception of the intercept estimator ($\hat{\alpha}$) in the AR-MAR model, in the AR-MAR model these estimators display low bias even in high noise conditions).

Table 6. Averaged bootstrapped standard error for the bias (IntAR-MAR model).

	$\hat{\alpha}$	$\hat{\beta}_\alpha$	$\hat{\rho}$	$\hat{\beta}_\rho$
100 Observations	0.0005	0.0008	0.0006	0.0007
200 Observations	0.0003	0.0005	0.0004	0.0005
500 Observations	0.0002	0.0004	0.0002	0.0003
Low ES	0.0003	0.0005	0.0005	0.0006
Medium ES	0.0003	0.0006	0.0004	0.0005
High ES	0.0004	0.0007	0.0004	0.0005
Type 2	0.0003	0.0006	0.0004	0.0005
Type 1	0.0004	0.0007	0.0005	0.0006
Type 3	0.0003	0.0006	0.0004	0.0005
Fast-switching	0.0003	0.0004	0.0003	0.0004
Slow-switching	0.0003	0.0005	0.0004	0.0005
0	0.0005	0.0009	0.0005	0.0007
0.13	0.0005	0.0008	0.0005	0.0007
0.23	0.0004	0.0007	0.0005	0.0007
0.31	0.0004	0.0007	0.0005	0.0007
0.43	0.0004	0.0007	0.0005	0.0007

In contrast to the bias, the main effect of the variance of the estimators are less clear cut. However, the following general main effects of estimator variance can be seen across the MAR models in Tables 1–3: (1) The variance of estimators consistently decreases with higher numbers of observations. (2) The variance of all estimators increases with the effect size of the covariate influence, the higher the covariate influence the higher the estimator variance. (3) Estimator variance is often lowest for the Constant continuous noise covariate. (4) In terms of estimator variance there is no clear main effect for the switching frequency of the covariate. (5) In terms of estimator variance there is no clear main effect for the amount of noise in the covariate. Below we discuss the implications of our simulation for study design and applied research.

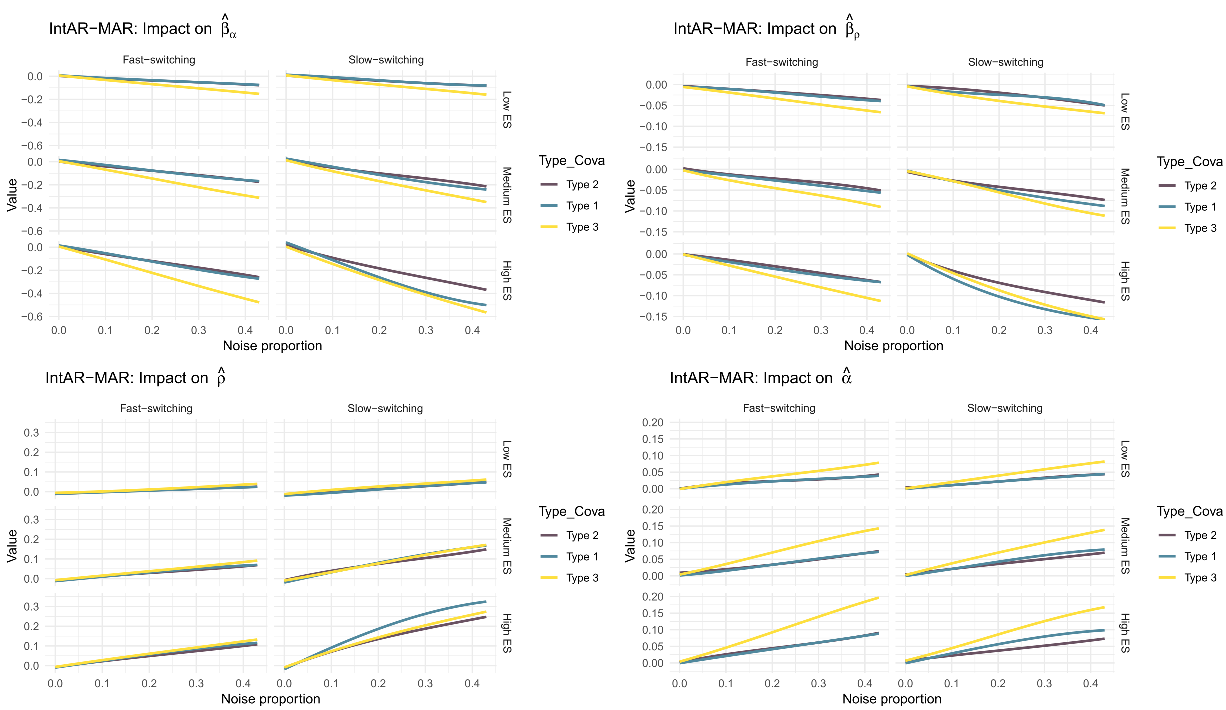

Figure 5. Average bias in estimators of the IntAR-MAR model as a function of covariate noise, under various conditions.
Note. Type 1: Occasional discrete noise, Type 2: Occasional continuous noise, Type 3: Constant continuous noise. Low ES: conditions where $\beta_\rho = 0.15$ and $\beta_\alpha = 0.36$, Medium ES: conditions where $\beta_\rho = 0.20$ and $\beta_\alpha = 0.72$, High ES: conditions where $\beta_\rho = 0.25$ and $\beta_\alpha = 1.08$.

To facilitate researchers in utilizing these simulation results for assessing the extent to which effect sizes are impacted by covariate noise, Figure 5 illustrates the average bias for the IntAR-MAR model as a function of covariate noise, across various data characteristics. Correspondingly, Figures S6 and S7 in the Supplemental material, supplemental content C show these average biases as a function of covariate noise, for the Int-MAR and the AR-MAR model, respectively. We created these figures by drawing a spline through the average bias that we observed for different proportions of noise in our simulation. These figures thus allow researchers to evaluate and compare different study designs. For example, when considering a study design with an IntAR-MAR model featuring medium effect sizes for β_ρ and β_α, along with a slow-switching, Constant continuous noise covariate, the top-right panel in Figure 5 illustrates that if this study design employs a covariate with a noise proportion of 0.35, the estimate $\hat{\beta}_\rho$ can be expected to decrease by 0.1 compared to the true parameter value (i.e., bias). However, if the noise proportion of the covariate is lowered to 0.1, Figure 5 indicates that the estimate $\hat{\beta}_\rho$ would only decrease by roughly 0.025.

Discussion

While the effects of measurement error (i.e., noise) have received ample attention for cross-sectional regression models (Wooldridge, 2002), in this paper we have investigated its effects for longitudinal regression models, like the MAR model, where their effects are not often considered (for an exception see Castro-Alvarez et al., 2022; Schuurman et al., 2015). Specifically, in this paper we have investigated different scenarios where the *covariate* is a noisy predictor of the changes in a process due to measurement error.

Our simulation showed that noise in the *covariate* will lead to bias in almost all estimators of the MAR model. This bias was already notable for a noise proportion of 13%. When the noise proportion was 43%, the bias was large for nearly all estimators: the AR estimator in the IntAR-MAR model, for instance, exhibited a relative bias equal to 60% of the true effect size. Our results are concerning because measuring a covariate with a noise proportion of 43% is plausible in psychological research practice (Schmidt & Hunter, 1996). This bias of estimators can have implications for the conclusions drawn from MAR models. Specifically, noise in the covariate attenuates estimates for the influence of the covariate toward zero (known as regression dilution bias (MacMahon et al., 1990)). Though our simulation covered only positive parameter values for these regression coefficients, noise in the covariate is known to bias the estimate of the associated regression coefficient toward zero, as long as only a single predictor variable is used and the noise is unbiased and independent of the true value of the outcome variable (Frost & Thompson, 2000). Further, noise in the covariate inflates estimates for the autoregression of the outcome variable. Thus, noise in the covariate leads to incorrect conclusions, suggesting less absolute influence of the covariate on the dynamics and a too large autoregression.

The aim of this paper is not to discourage people from including contextual influences into the AR model. On the contrary, if a covariate indeed affects the emotion dynamics under study, omitting it will lead to biased estimates. Instead, we advise researchers to adjust their study design accordingly and to consider the accuracy with which a covariate can be measured when interpreting the results of a MAR model. This is particularly important because recent research suggests that applied researchers often do not consider the psychometric properties of their intensive longitudinal data (Vogelsmeier et al., 2024). The remainder of this paper will be dedicated to a number of implications of our simulation findings for applied research and a discussion on possible solutions to address the problem of noisy covariates in MAR models. We will conclude with a few general considerations for specifying a MAR model.

Implications for applied research

Our simulation showed that when a noisy covariate persists over time (i.e., switches slowly), this can particularly inflate estimates for the AR coefficient (all covariates in our simulation design persisted for some observations, though the slow-switching covariates persisted longer than the fast-switching covariates). Thus, the higher the temporal persistence of a noisy covariate, the higher the positive bias of the AR coefficient. This finding has important implications because the AR coefficient is often of particular interest to psychological researchers (see e.g., Brose et al., 2015b; Koval & Kuppens, 2012; Koval et al., 2012; Kuppens et al., 2010) and because many of the contextual covariates that are relevant in psychology show some temporal persistence (e.g., work demands, or weather often persist for several observations).[17] Thus, if permitted by the research question, studies should favor designs where the

[17]Though contextual covariates without temporal persistence also occur in empirical studies, for instance when including a covariate that indicates whether a random prompt during an ecological momentary assessment was made in the morning or in the evening.

covariate exhibits low temporal persistence (i.e., fast-switching). Otherwise, the temporal persistence in a noisy covariate can be mistaken for higher autoregression in the outcome variable, as was shown in our simulation. Our advice is in line with the advice given by Ariens et al. (2023) who studied noise-free MAR models and noted that serial dependence in the covariate can lead to lower estimation accuracy.

In our simulation, the Occasional continuous noise covariate was associated with the lowest bias, followed by the Occasional discrete noise covariate, the Constant continuous noise covariate had the highest bias. This suggest firstly that in the presence of comparable proportions of noise, continuous covariates with occasional noise display lower bias than discrete covariates with occasional noise. Secondly, in the presence of comparable proportions of noise, occasional noise is associated with lower bias than constant noise. Thus, if permitted by the covariate under study, researchers should favor continuous rather than discrete covariates. These findings suggest also that researchers should take steps to avoid study designs where all observations of the covariate are noisy. For instance, by avoiding a study design that includes measurement times at which the participant is particularly prone to enter inaccurate responses due to being in-attentive, tired, or unable to take enough time to enter responses. Generally, keeping the total amount of noise to a minimum is crucial because the bias that is caused by a noisy covariate quickly increased as the amount of noise in the covariate increased, as was shown in our simulation study. Further, the bias caused by noisy covariates increased as the effect size of the covariate increased. Consequently, researchers should particularly avoid noise in the measurement of the covariate when the covariate is expected to have a large effect on the emotion dynamics. The bias that results from a noisy covariate did not decrease as the number of observations increased. Hence, researchers should be mindful of the bias that is caused by a noisy covariate even when they collect very long time-series.

Finally, when considering tradeoffs in the study design and analysis of intensive longitudinal data, researchers can utilize Figures S5–S7 (Figures S6 and S7 are contained in Supplemental material, supplemental content C). These figures show how the parameter estimates of different MAR models are impacted as a function of covariate noise, across various data characteristics. By locating the appropriate panel and curve given the characteristics of the data that is being collected, researchers can evaluate the anticipated changes in effect size due to noise in the covariate. These figures provide a way for researchers to determine whether they deem the degree of bias that is caused by a given noise proportion as acceptable or whether efforts need to be made to collect a less noisy covariate.

Statistical solutions

Schuurman et al. (2015) investigated the consequences of measurement error in the outcome variable in AR models, they showed that the bias resulting from this can be alleviated by explicitly modeling the outcome variable's measurement error. Though Schuurman et al. (2015) studied only the scenario of noise in the outcome variable and not of noise in the covariate, an equivalent approach (i.e., explicitly modeling the covariate's measurement error) could be used to counteract the bias we have found in our simulation. Explicitly modeling measurement error, however, relies on assumptions, such as a Normal distribution of the measurement error. Thus, explicitly modeling measurement error will make a statistical model more complex and harder to estimate. In a Bayesian framework, for instance, estimating a MAR model that accounts for measurement error of the covariate will require strong informative prior distributions, especially for short time-series.

Various methods could be used to estimate a MAR model where measurement error in the covariate is accounted for through a measurement model. For instance, by using the dynamic factor analysis model offered in Mplus (Asparouhov et al., 2018; Muthén & Muthén, 2013), or by using state space model estimation as offered in the dynr R package (Ou et al., 2019). Measurement error could also be incorporated into a MAR model through structural equation modeling (Bollen, 1989; Jaccard & Wan, 1995). Alternatively, the development of errors-in-variables methods for MAR models could prove beneficial to achieve consistent estimates despite measurement error, for instance through instrumental variable methods which are distribution-free and have a certain robustness to model misspecification (Bollen et al., 2007, 2024). Instrumental variable methods use so-called 'instruments': variables that are uncorrelated with the measurement error but are correlated with the covariate. Once instrumental variables are identified, instrumental variable estimators can be used to estimate the latent variable model, for instance in a factor model (Bollen, 1996).

While such methods to address measurement error are already widely developed for cross-sectional and

AR models (Castro-Alvarez et al., 2022; Schuurman et al., 2015), their estimation possibilities for *moderated* AR models, like the MAR, are still lagging behind. Studying potential modeling solutions to account for noisy covariates in MAR models would be very beneficial for intensive longitudinal studies of psychological concepts.

Relation to omitted variables and other sources of model misspecification

In this paper we have focused exclusively on noise being due to measurement error. Noise can also arise, however, when there are omitted variables. That is, when there are unobserved covariates that correlate with the observed predictor and that also cause changes in the emotion dynamics. While omitted variables and measurement error are conceptually very different, they are similar in their statistical structure (see e.g., Wooldridge, 2002, p. 70): In both cases the covariate is an imperfect predictor of the changes in the dynamics and thus an unexplained part (unexplained either because the variable is omitted or because it is measurement error) ends up in the innovation. In the measurement error and the omitted variable case, this can lead to a correlation between predictor x_{t-1} and innovation ζ_t (see Wooldridge, 2002, pp. 71–76) or our illustration in the Supplemental material, supplemental content A), violating thus the assumption of predictors being uncorrelated with the innovation. In short, while our simulation included only the measurement error scenario, it also covers a special case of the omitted variable problem. Thus, similar bias than we found in our simulation can be expected when important predictors that covary with the included covariate are omitted form the MAR model.

In addition to measurement error and omitted variables, other sources of model misspecification are possible. For instance, it is a crucial consideration whether to model the covariate's influence on an intercept (as specified in Equations (2)–(4)) or on a within-person mean (as specified in Equations (A.9)–(A.12) in Supplemental material, supplemental content D). This consideration is essential because the interpretation of the covariate influence differs between these two model specifications, and furthermore, these specifications imply different assumptions about the underlying mechanism generating the data (Ernst et al., 2024; Usami et al., 2019). In the intercept specification, the influence of the covariate is assumed to carry over to the next time-points *via* the autoregressive process. In contrast, in the within-person mean specification, the covariate is assumed to influence only one measurement, and these influences do not carry over to any future time-point. Consequently, covariates included in a model with intercept-specification are commonly referred to as accumulating factors (Ernst et al., 2024; Usami et al., 2019), while covariates included in a model with within-person mean specification are referred to as deterministic trends (Usami et al., 2019) or as direct influences (Ernst et al., 2024). For details on the distinction and the implications of these two specifications see Ernst et al. (2024), Hamaker (2005), and Usami et al. (2019).

Thus, because these two specifications are not equivalent, the model is misspecified when the data is generated according to an intercept specification but is analyzed with a within-person mean specification, or vice versa. We expect that such misspecification will cause similar bias as when a misspecification occurs because of measurement error or omitted variables. In the future, it would be useful to investigate and compare the amount of bias that can result from these different model misspecifications.

In this paper we have considered exclusively the scenario where data is generated and analyzed according to the intercept specification, in order to keep the focus of our paper on misspecification due to noise in the covariate. In the Supplemental material, supplemental content D we show the results of an additional simulation where data was generated and analyzed according to a MAR model with a within-person mean specification. The results of this simulation are presented in Table 7 and Figure S8 in the Supplemental material, supplemental content D. These results show that the main effects for the bias in MAR models with within-person mean specification are the same as those for MAR models with intercept specification (the effects for the different types of covariate, however, were less clear-cut in this simulation). Also, equivalently to MAR models with intercept specification, for MAR models with within-person mean specification, noise in the covariate attenuates estimates for the influence of the covariate toward zero while inflating estimates for the autoregression.

Extensions to multi-individual MAR models

In this paper we have focused on single-individual models rather than multi-individual models. We did so because for multi-individual models there is an added layer of complexity and within-person and between-person parameter estimators become related and the

Table 7. Main effects for bias and variance across the simulation factors (WpmAR-MAR model).

	$\hat{\alpha}$ Bias	$\hat{\alpha}$ Var	$\hat{\beta}_\alpha$ Bias	$\hat{\beta}_\alpha$ Var	$\hat{\rho}$ Bias	$\hat{\rho}$ Var	$\hat{\beta}_\rho$ Bias	$\hat{\beta}_\rho$ Var
100 Observations	.077	.023	−.175	.099	.001	.021	−.048	.035
200 Observations	.076	.013	−.175	.080	.013	.010	−.042	.017
500 Observations	.076	.007	−.177	.070	.019	.004	−.039	.008
Low ES	.038	.010	−.082	.018	.001	.011	−.028	.018
Medium ES	.076	.012	−.171	.031	.009	.011	−.041	.019
High ES	.116	.018	−.274	.061	.023	.014	−.060	.021
Type 2	.060	.012	−.138	.083	.005	.011	−.038	.020
Type 1	.076	.016	−.182	.085	.015	.014	−.032	.025
Type 3	.094	.015	−.207	.079	.014	.011	−.059	.016
Fast-switching	.068	.012	−.149	.081	.002	.011	−.043	.019
Slow-switching	.085	.017	−.202	.084	.020	.013	−.042	.021
0	$\lvert x\rvert < .001$	.009	.001	.105	−.015	.011	−.006	.020
0.13	.046	.010	−.111	.079	−.006	.011	−.029	.019
0.23	.082	.011	−.194	.065	.010	.012	−.046	.020
0.31	.108	.013	−.248	.056	.021	.012	−.057	.019
0.43	.145	.016	−.326	.046	.045	.013	−.077	.019

true parameter values at both levels, the number of observations, and the number of persons will have an impact on estimator performance at both levels, which would make our results even harder to communicate (Schultzberg & Muthén, 2018). However, we expect multi-individual extensions of the MAR model to show similar biases overall due to noisy covariates as we have shown for single-individual models. We expect this because we have shown in our simulation that these biases do not decrease with an increase in available data (i.e., by increasing the number of observations). Multi-individual extensions of MAR models are possible, for instance through multilevel extensions (Ernst et al., 2021). Multilevel MAR models can be estimated in MPlus (Muthén & Muthén, 2013)[18] while latent class MAR models can be estimated through ClusterVAR in R (Ernst & Haslbeck, 2024).[19]

Conclusion

In sum, while it is crucial for researchers to account for contextual events when modeling emotion dynamics, researchers should be mindful of the noise that is contained in measures of such contextual events because this noise can cause large bias of parameter estimates. The bias caused by noisy measures of contextual events is further exacerbated by temporal persistence in the contextual events, discrete measures of the contextual events, a larger effect of the contextual events, and by a constant rather than occasional presence of noise. The bias that results from noisy measures of contextual events does not decrease as the number of observations increases.

[18]MPlus employs only the within-person mean specification, not the intercept specification. Also, to estimate an AR-MAR model in Mplus requires the use of a work-around, see Koval and Kuppens (2012).

[19]ClusterVAR can only estimate the within-person mean specification and does not facilitate the estimation of AR-MAR.

Article information

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

Funding: The research presented in this article was supported by research grants from the Fund for Scientific Research-Flanders (FWO; EOS40007528/G0I2422N) and from the Research Council of KU Leuven (C14/23/062; iBOF/21/090) awarded to E. Ceulemans.

Data availability statement: The R-code to replicate all simulations and analysis that are presented in this manuscript are published on OSF. No empirical data is used in this manuscript.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design

Author contributions: AFE, EC, and JA conceived the presented ideas. AFE performed the simulation and analyzed the results. JA wrote the simulation code for the data generation. JA made Figures 1 and 2. AFE wrote the manuscript with the input from all authors. All authors contributed to the final manuscript.

References

Adolf, J. (2023). fmTSA. https://gitlab.kuleuven.be/ppw-okpiv/researchers/u0119417/published/fmTSA

Adolf, J. K., Voelkle, M. C., Brose, A., & Schmiedek, F. (2017). Capturing context-related change in emotional dynamics via fixed moderated time series analysis. *Multivariate Behavioral Research*, *52*(4), 499–531. https://doi.org/10.1080/00273171.2017.1321978

Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management*, *21*(6), 1141–1158. https://doi.org/10.1016/0149-2063(95)90026-8

Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology*, *82*(1), 192–206. http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true&db=psyh&AN=1997-07782-013&site=ehost-live&scope=site https://doi.org/10.1037/0021-9010.82.1.192

Albers, C. J., & Bringmann, L. F. (2020). Inspecting gradual and abrupt changes in emotion dynamics with the time-varying change point autoregressivemodel. *European Journal of Psychological Assessment*, *36*(3), 492–499. https://doi.org/10.1027/1015-5759/a000589

Ariens, S., Adolf, J. K., & Ceulemans, E. (2023). Collinearity issues in autoregressive models with time-varying serially dependent covariates. *Multivariate Behavioral Research*, *58*(4), 687–705. https://doi.org/10.1080/00273171.2022.2095247

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons. https://doi.org/10.1002/9781118619179

Bollen, K. A. (1996). An alternative two stage least squares (2sls) estimator for latent variable equations. *Psychometrika*, *61*(1), 109–121. https://doi.org/10.1007/BF02296961

Bollen, K. A., Gates, K. M., & Luo, L. (2024). A model implied instrumental variable approach to exploratory factor analysis (MIIV-EFA). *Psychometrika*, *89*(2), 687–716. https://doi.org/10.1007/s11336-024-09949-6

Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., & Chen, F. (2007). Latent variable models under misspecification: Two-stage least squares (2sls) and maximum likelihood (ml) estimators. *Sociological Methods & Research*, *36*(1), 48–86. https://doi.org/10.1177/0049124107301947

Bringmann, L. F., Ariens, S., Ernst, A. F., Snippe, E., & Ceulemans, E. (2024). Changing networks: Moderated idiographic psychological networks. *Advances in Psychology*, *2*, e658296. https://doi.org/10.56296/aip00014

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling nonstationary emotion dynamics in dyads using a time-varying vector autoregressive model. *Multivariate Behavioral Research*, *53*(3), 293–314. https://doi.org/10.1080/00273171.2018.1439722

Brose, A., Schmiedek, F., Koval, P., & Kuppens, P. (2015a). Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition & Emotion*, *29*(3), 527–538. https://doi.org/10.1080/02699931.2014.916252

Brose, A., Voelkle, M. C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2015b). Differences in the between-person and within-person structures of affect are a matter of degree. *European Journal of Personality*, *29*(1), 55–71. https://doi.org/10.1002/per.1961

Cabrieto, J., Adolf, J., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2018). Detecting long lived autodependency changes in a multivariate system via change point detection and regime switching models. *Scientific Reports*, *8*(1), 15637. https://doi.org/10.1038/s41598-018-33819-8

Castro-Alvarez, S., Tendeiro, J. N., de Jonge, P., Meijer, R. R., & Bringmann, L. F. (2022). Mixed-effects trait-state-occasion model: Studying the psychometric properties and the person–situation interactions of psychological dynamics. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(3), 438–451. https://doi.org/10.1080/10705511.2021.1961587

Chow, S.-M., Grimm, K. J., Filteau, G., Dolan, C. V., & McArdle, J. J. (2013). Regime switching bivariate dual change score model. *Multivariate Behavioral Research*, *48*(4), 463–502. https://doi.org/10.1080/00273171.2013.787870

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/-correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates Publishers.

Crayen, C., Eid, M., Lischetzke, T., & Vermunt, J. K. (2017). A continuous-time mixture latent-state-trait Markov model for experience sampling data: Application and evaluation. *European Journal of Psychological Assessment*, *33*(4), 296–311. https://doi.org/10.1027/1015-5759/a000418

Curran, P. J., & Bauer, D. J. (2007). Building path diagrams for multilevel models. *Psychological Methods*, *12*(3), 283–297. https://doi.org/10.1037/1082-989X.12.3.283

Dejonckheere, E., Mestdagh, M., Kuppens, P., & Tuerlinckx, F. (2020). Reply to: Context matters for affective chronometry. *Nature Human Behaviour*, *4*(7), 690–693. https://doi.org/10.1038/s41562-020-0861-6

Dunlap, W. P., & Kemery, E. R. (1988). Effects of predictor intercorrelations and reliabilities on moderated multiple regression. *Organizational Behavior and Human Decision Processes*, *41*(2), 248–258. https://doi.org/10.1016/0749-5978(88)90029-5

Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica*, *51*(2), 277–304. https://doi.org/10.2307/1911990

Ernst, A. F., Albers, C. J., & Timmerman, M. E. (2024). A comprehensive model framework for between-individual differences in longitudinal data. *Psychological Methods*, *29*(4), 748–766. https://doi.org/10.1037/met0000585

Ernst, A. F., Albers, C. J., Jeronimus, B. F., & Timmerman, M. E. (2020). Inter-individual differences in multivariate time-series: Latent class vector-autoregressive modeling. *European Journal of Psychological Assessment*, *36*(3), 482–491. https://doi.org/10.1027/1015-5759/a000578

Ernst, A. F., Haslbeck, J. M. B. (2024). *Clustervar: Fitting latent class vector autoregressive (var) models [R package version 0.0.7]*. https://cran. r-project.org/web/packages/ClusterVAR/index.html

Ernst, A. F., Timmerman, M. E., Jeronimus, B. F., & Albers, C. J. (2021). Insight into individual differences in emotion dynamics with clustering. *Assessment*, *28*(4), 1186–1206. https://doi.org/10.1177/1073191119873714

Frost, C., & Thompson, S. G. (2000). Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *163*(2), 173–189. https://doi.org/10.1111/1467-985X.00164

Fuchs, P., Nussbeck, F. W., Meuwly, N., & Bodenmann, G. (2017). Analyzing dyadic sequence data—Research questions and implied statistical models. *Frontiers in Psychology*, *8*, 429. https://doi.org/10.3389/fpsyg.2017.00429

Griffin, W. A., & Li, X. (2016). Using Bayesian nonparametric hidden semi-Markov models to disentangle affect processes during marital interaction. *PloS One*, *11*(5), e0155706. https://doi.org/10.1371/journal.pone.0155706

Haan-Rietdijk, S., Gottman, J. M., Bergeman, C. S., & Hamaker, E. L. (2016). Get over it! A multilevel threshold autoregressive model for state-dependent affect regulation. *Psychometrika*, *81*(1), 217–241. https://doi.org/10.1007/s11336-014-9417-x

Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research*, *33*(3), 404–416. https://doi.org/10.1177/0049124104270220

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the cogito study. *Multivariate Behavioral Research*, *53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Hamaker, E. L., Grasman, R. P. P. P., & Kamphuis, J. H. (2016). Modeling BAS dysregulation in bipolar disorder: Illustrating the potential of time series analysis. *Assessment*, *23*(4), 436–446. https://doi.org/10.1177/1073191116632339

Hamilton, J. (1994). *Time series analysis*. Princeton University Press.

Haslbeck, J. M. B., Borsboom, D., & Waldorp, L. J. (2021). Moderated network models. *Multivariate Behavioral Research*, *56*(2), 256–287. https://doi.org/10.1080/00273171.2019.1677207

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer. https://books.google.nl/books?id=eBSgoAEACAAJ

Hendry, D. F. (1995). *Dynamic econometrics*. The Clarendon Press, Oxford University Press. https://doi.org/10.1093/0198283164.001.0001

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*(3), 329–367. https://doi.org/10.1177/0049124198026003003

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*(4), 901–930. https://doi.org/10.1037/a0038822

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, *117*(2), 348–357. https://doi.org/10.1037/0033-2909.117.2.348

Jongerling, J., Laurenceau, J.-P., & Hamaker, E. L. (2015). A multilevel AR(1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research*, *50*(3), 334–349. https://doi.org/10.1080/00273171.2014.1003772

Kang, S.-M., & Waller, N. G. (2005). Moderated multiple regression, spurious interaction effects, and IRT. *Applied Psychological Measurement*, *29*(2), 87–105. https://doi.org/10.1177/0146621604272737

Koval, P., & Kuppens, P. (2012). Changing emotion dynamics: Individual differences in the effect of anticipatory social stress on emotional inertia. *Emotion (Washington, D.C.)*, *12*(2), 256–267. https://doi.org/10.1037/a0024756

Koval, P., Kuppens, P., Allen, N. B., & Sheeber, L. (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition & Emotion*, *26*(8), 1412–1427. https://doi.org/10.1080/02699931.2012.667392

Kuppens, P., & Verduyn, P. (2017). Emotion dynamics. *Current Opinion in Psychology*, *17*, 22–26. https://doi.org/10.1016/j.copsyc.2017.06.004

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*(7), 984–991. https://doi.org/10.1177/0956797610372634

Lancee, J., Harvey, A., Morin, C., Ivers, H., van der Zweerde, T., & Blanken, T. (2022). Network intervention analyses of cognitive therapy and behavior therapy for insomnia: Symptom specific effects and process measures. *Behaviour Research and Therapy*, *153*, 104100. https://doi.org/10.1016/j.brat.2022.104100

Lapate, R. C., & Heller, A. S. (2020). Context matters for affective chronometry. *Nature Human Behaviour*, *4*(7), 688–689. https://doi.org/10.1038/s41562-020-0860-7

Liu, S., Ou, L., & Ferrer, E. (2021). Dynamic mixture modeling with dynr. *Multivariate Behavioral Research*, *56*(6), 941–955. https://doi.org/10.1080/00273171.2020.1794775

Liu, Y., & Salvendy, G. (2009). Effects of measurement errors on psychometric measurements in ergonomics studies: Implications for correlations, ANOVA, linear regression, factor analysis, and linear discriminant analysis. *Ergonomics*, *52*(5), 499–511. https://doi.org/10.1080/00140130802392999

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

MacMahon, S., Peto, R., Cutler, J., Collins, R., Sorlie, P., Neaton, J., Abbott, R., Godwin, J., Dyer, A., & Stamler, J. (1990). Blood pressure, stroke, and coronary heart disease. Part 1, prolonged differences in blood pressure: Prospective observational studies corrected for the regression dilution bias. *Lancet (London, England)*, *335*(8692), 765–774. https://doi.org/10.1016/0140-6736(90)90878-9

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, *25*(5), 610–635. https://doi.org/10.1037/met0000250

Mestdagh, M., & Dejonckheere, E. (2021). Ambulatory assessment in psychopathology research: Current achievements and future ambitions. *Current Opinion in Psychology*, *41*, 1–8. https://doi.org/10.1016/j.copsyc.2021.01.004

Muthén, L. K., & Muthén, B. O. (2013). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling.

Psychometrika, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8
Ou, L., Hunter, M. D., & Chow, S.-M. (2019). Whats for dynr: A package for linear and nonlinear dynamic modeling in R. *R Journal*, *11*(1), 91–111. https://doi.org/10.32614/RJ-2019-012
Pesaran, M. H. (2015). *Time series and panel data econometrics*. Oxford. https://doi.org/10.1093/acprof:Oso/9780198736912.001.0001
Pfaff, B. (2008). Var, svar and svec models: Implementation within R package vars. *Journal of Statistical Software*, *27*(4), 1–32. https://www.jstatsoft.org/v27/i04/ https://doi.org/10.18637/jss.v027.i04
Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*(2), 199–223. https://doi.org/10.1037/1082-989X.1.2.199
Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 495–515. https://doi.org/10.1080/10705511.2017.1392862
Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology*, *6*, 1038. https://doi.org/10.3389/fpsyg.2015.01038
Sels, L., Cabrieto, J., Butler, E., Reis, H., Ceulemans, E., & Kuppens, P. (2020). The occurrence and correlates of emotional interdependence in romantic relationships. *Journal of Personality and Social Psychology*, *119*(1), 136–158. https://doi.org/10.1037/pspi0000212
Sels, L., Schat, E., Verhofstadt, L., & Ceulemans, E. (2022). Introducing change point detection analysis in relationship research: An investigation of couples' emotion dynamics. *Journal of Social and Personal Relationships*, *39*(10), 3133–3154.
Simons, J. S., Simons, R. M., Grimm, K. J., Keith, J. A., & Stoltenberg, S. F. (2021). Affective dynamics among veterans: Associations with distress tolerance and post traumatic stress symptoms. *Emotion (Washington, D.C.)*, *21*(4), 757–771. https://doi.org/10.1037/emo0000745
Speyer, L. G., Murray, A. L., & Kievit, R. (2024). Investigating moderation effects at the within-person level using intensive longitudinal data: A two-level dynamic structural equation modelling approach in mplus. *Multivariate Behavioral Research*, *59*(3), 620–637. https://doi.org/10.1080/00273171.2023.2288575
Stan Development Team. (2024). *RStan: The R interface to Stan.* [R package version 2.32.5]. https://mc-stan.org/
Stifter, C. A., & Rovine, M. (2015). Modeling dyadic processes using hidden Markov models: A time series approach to mother-infant interactions during infant immunization. *Infant and Child Development*, *24*(3), 298–321. https://doi.org/10.1002/icd.1907
Stone-Romero, E. F., & Anderson, L. E. (1994). Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology*, *79*(3), 354–359. https://doi.org/10.1037/0021-9010.79.3.354
Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657. https://doi.org/10.1037/met0000210
van Roekel, E., Verhagen, M., Engels, R. C. M. E., & Kuppens, P. (2018). Variation in the serotonin transporter polymorphism (5-httlpr) and inertia of negative and positive emotions in daily life. *Emotion (Washington, D.C.)*, *18*(2), 229–236. https://doi.org/10.1037/emo0000336
Vogelsmeier, L. V. D. E., Jongerling, J., & Maassen, E. (2024). Assessing and accounting for measurement in intensive longitudinal studies: Current practices, considerations, and avenues for improvement. *Quality of Life Research*, *33*(8), 2107–2118. https://doi.org/10.1007/s11136-024-03678-0
Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.
Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach*. South-Western. http://books.google.ch/books?id=64vt5TDBNLwC