

ZERO-INFLATED REGIME-SWITCHING STOCHASTIC DIFFERENTIAL EQUATION
MODELS FOR HIGHLY UNBALANCED MULTIVARIATE, MULTI-SUBJECT
TIME-SERIES DATA

ZHAO- HUA LU 

ST. JUDE CHILDREN'S RESEARCH HOSPITAL

SY- MIIN CHOW, NILAM RAM AND PAMELA M. COLE

THE PENNSYLVANIA STATE UNIVERSITY

In the study of human dynamics, the behavior under study is often operationalized by tallying the frequencies and intensities of a collection of lower-order processes. For instance, the higher-order construct of negative affect may be indicated by the occurrence of crying, frowning, and other verbal and nonverbal expressions of distress, fear, anger, and other negative feelings. However, because of idiosyncratic differences in how negative affect is expressed, some of the lower-order processes may be characterized by sparse occurrences in some individuals. To aid the recovery of the true dynamics of a system in cases where there may be an inflation of such “zero responses,” we propose adding a regime (unobserved phase) of “non-occurrence” to a bivariate Ornstein–Uhlenbeck (OU) model to account for the high instances of non-occurrence in some individuals while simultaneously allowing for multivariate dynamic representation of the processes of interest under nonzero responses. The transition between the occurrence (i.e., active) and non-occurrence (i.e., inactive) regimes is represented using a novel latent Markovian transition model with dependencies on latent variables and person-specific covariates to account for inter-individual heterogeneity of the processes. Bayesian estimation and inference are based on Markov chain Monte Carlo algorithms implemented using the JAGS software. We demonstrate the utility of the proposed zero-inflated regime-switching OU model to a study of young children's self-regulation at 36 and 48 months.

Key words: stochastic differential equations, Ornstein–Uhlenbeck, Markov switching transition, regime switching, Bayesian methods, Markov chain Monte Carlo algorithms.

The past decade has evidenced tremendous growth in the development and application of differential equation models as a representation of change processes in the social and behavioral sciences. The need to apply and develop more sophisticated methods for representing change is instigated in part by the growing prevalence of intensive longitudinal data (ILD) such as physiological data (Wilhelm, Grossman, & Muller, 2012; M. Yang & Chow, 2010) and brain imaging data (Gates & Molenaar, 2012). Stochastic differential equation (SDE) models have gained popularity in the psychometric literature as a way to analyze ILD, either in the form of linear differential equation models (Arminger, 1986; Coleman, 1968; Oravecz, Tuerlinckx, & Vandekerckhove, 2011; Oud & Jansen, 2000; Oud & Singer, 2008; Singer, 2010, 2012; Voelkle, Oud, Davidov, & Schmidt, 2012), or nonlinear differential equation models (Lu, Chow, Sherwood, & Zhu, 2015; Molenaar & Newell, 2003; Singer, 1992, 2010, 2012). Ordinary differential equations characterize the underlying mechanisms of dynamic processes through explicit specifications of the relations between the dynamic processes of interest, and their derivatives (i.e., instantaneous

Zhao-Hua Lu and Sy-Miin Chow have contributed equally to this work.

Funding for this study was provided by NSF Grant SES-1357666, NIH Grants R01MH61388, R01HD07699, R01GM105004, U24EB026436, Penn State Quantitative Social Sciences Initiative and UL TR000127 from the National Center for Advancing Translational Sciences. The article is partly done when Zhao-Hua Lu was in the Pennsylvania State University.

Correspondence should be made to Zhao-Hua Lu, St. Jude Children's Research Hospital, MS 768, Room R6006, 262 Danny Thomas Place, Memphis, TN 38105-3678, USA. Email: zhaohua.lu@stjude.org

changes and other higher-order changes therein). SDEs incorporate additional stochastic process noises into ordinary differential equations to account for random fluctuations in those processes. Additionally, differential equations can readily accommodate irregular intervals between successive measurements and are thus especially conducive as a tool in other applications involving irregularly spaced ILD.

The key challenge we seek to address in the present article is to find ways to meaningfully represent the dynamics of certain behaviors/emotions measured by ILD when there is sparseness in particular response categories. Due to individual differences, varying target behaviors across task time and many other reasons, some behaviors may be characterized by high instances of non-occurrence (coded as zero)—in other words, inflation in zero responses. Zero-inflated models have been studied in cross-sectional data (Lambert, 1992) and traditional longitudinal data (Hall, 2000), but less so in the context of ILD. Our motivating example features one such examples from a emotion regulation study involving young children, in which zero inflation is present due in part to developmental reasons. Similar instances of zero inflation have also been observed in other intensive longitudinal laboratory data (e.g., facial electromyography data; M. Yang & Chow, 2010), as well as substance use data following treatment of alcohol use disorder (Maisto et al., 2017). Differential equation models provide a way to study relations between intensive moment-to-moment dynamics and test whether specific strategies influence change in children's negative emotion

High instances of non-occurrence in ILD pose various estimation challenges. The dynamic mechanisms of the non-occurrence and occurrence periods are usually fundamentally different. Using one single SDE model to represent such distinct dynamics can be challenging and, in some cases, would yield biased estimates and interpretations of the dynamic system as a whole. To accommodate high instances of non-occurrence in ILD, we propose to include a zero inflation (ZI) component in a mixture SDE framework with regime switching to accomplish simultaneous representation of ZI and the dynamics of the system under nonzero responses.

Our operating dynamic model for portions of the data with nonzero responses assumes the form of an SDE model, specifically the Ornstein–Uhlenbeck (OU) model—a popular dynamic modeling framework in the econometric, engineering, and statistical literature widely used to characterize stochastic processes that fluctuate around an equilibrium (Ait-Sahalia, 2008; Beaulieu, Jhwueng, Boettiger, & O'Meara, 2012; Beskos, Papaspiliopoulos, & Roberts, 2009; Beskos, Papaspiliopoulos, Roberts, & Fearnhead, 2006; Jones, 1984; Mbalawata, Särkkä, & Haario, 2013; Ramsay, Hooker, Campbell, & Cao, 2007; Särkkä, 2013; Uhlenbeck & Ornstein, 1930). It has been used, for instance, as a model to represent individuals' emotion regulation (Oravec et al., 2011) and ambulatory blood pressure dynamics (Lu et al., 2015) due to its ability to capture—within a particular range of parameter values—homeostatic dynamics as exponential return to a baseline. This property of the OU model renders it especially appealing as a working model for the occurrence proportion of the ILD in our motivating example—a study examining developmental changes in children's self-regulation dynamics. However, a single OU model cannot be used to characterize the entire processes with occurrence and non-occurrence periods because the equilibria of the two periods can be very different.

Our work is unique and novel in a number of ways. First, our proposed model extends the classical OU model by allowing selected parameters from the model to differ depending on the latent phase—or regime—in which the process resides. The indicators of regime affiliation for different subjects and at different time points are similar to latent classes in mixture models. In particular, like latent classes, they are also unobserved latent variables whose actual values are unknown. However, unlike latent class models that assume class membership is a time-invariant characteristic of the person, regime-switching models allow individuals to switch between regimes over time as they transition through different phases of the change process (Kim & Nelson, 1999). The resulting regime-switching SDE modeling framework is distinct from conventional hidden

Markov models (Elliott, Aggoun, & Moore, 1995), or the related latent transition models (Collins & Wugalter, 1992; Lanza & Collins, 2008; Nylund, Muthén, Nishina, Bellmore, & Graham, 2006) because the SDEs allow the observed processes to evolve over time continuously.

Second, the proposed model allows greater flexibility than other regime-switching discrete-time dynamic models (e.g., Chow, Witkiewitz, Grasman, & Maisto, 2015; Chow & Zhang, 2013; Kim & Nelson, 1999; M. Yang & Chow, 2010) by allowing the dynamic processes to be defined in continuous time. Third, previous applications of regime-switching models in psychometrics have been restricted to models with a single-regime indicator (Chow, Grimm, Guillaume, Dolan, & McArdle, 2013; Chow & Zhang, 2013; Dolan, Schmittmann, Lubke, & Neale, 2005). However, when multivariate processes are involved and the timing at which each individual process transitions into and out of the ZI regime is disparate, this calls for the need to incorporate more than one-regime indicator, as is done in the present study. Along a similar line, our motivating example presents a novel demonstration that the interdependence of regime-switching between the two processes is, in and of itself, a question of substantive interest. The inclusion of covariates in the latent regime transition model further allows us to test postulates of age-related differences in such regime-switching dependencies. Fourth, the present study is the first at presenting a Bayesian framework for fitting an SDE model with regime-switching properties. Regime-switching OU models have been proposed in other fields (see, for example, Bai & Wu, 2018; J.-W. Yang, Tsai, Shyu, & Chang, 2016). However, they are all univariate models and cannot adequately characterize the third set of features described above.

The rest of the article is organized as follows. We first introduce a set of ILD from a self-regulation study with extended periods of consecutive zero responses, which motivated our proposed SDE model. Then, we propose a zero-inflated OU model (ZI-OU) and elucidate how it addresses the data analytic challenges described in the motivating example. We then outline the broader regime-switching SDE modeling framework within which the proposed ZI-OU model can be regarded as a special case. Next, we summarize the Bayesian estimation details and inference for this model, followed by results from fitting the proposed ZI-OU model to the empirical data, and a simulation study that serves to validate the targeted aspects of the estimation procedures. The paper is concluded with a discussion of the potential strengths and limitations of the proposed approach.

1. Motivating Example

Studies of young children's self-regulation are typically based on laboratory observations in which children are required to modulate reactions to task conditions that challenge their self-regulation. Self-regulation can be conceptualized as a multivariate process through which individuals engage in EP to delay, minimize, or desist PR (Baumeister & Vohs, 2007; Carver & Scheier, 1998; Kopp, 1982). PR are automatic reactions that are either learned or biologically prepared, and EP are actions involving higher-order psychological processes such as cognition and language.

The task used to elicit both anger and children's strategies for modulating anger in the current study is the transparent locked box task (Goldsmith & Reilly, 1993) designed to elicit negative emotions. In this task, each child chooses a desirable that is then locked in a clear acrylic box. The child is taught to open the box with a key, but is left alone with the box and the wrong set of keys. Effective self-regulation in this task requires the child to persist at opening the locked box despite intermittent manifestation of multiple negative emotions (e.g., anger, sadness) and off-task behaviors (e.g., pleading to the mother for help; engaging in other forms of distractions). As in many other studies that utilized the lock box task, researchers in our motivating study video-recorded children in the laboratory during the task and coded the presence and absence of markers of EP and PR second-by-second into a set of multi-subject, multivariate binary ILD (Cole et al.,

2011). Here, the sum of EP and PR marker scores (e.g., crying, expression of anger) per second is used as the EP and PR scores, respectively.

Plots of the moving averages with a window of ± 3 s of EP and PR scores from two randomly selected children at two ages are shown in Fig. 1. A few features of the data can be noted from the plots. First, even though composite EP and PR scores are being plotted, there is still considerable “sparseness” in the data, corresponding to periods of time in which none of the EP or PR markers were observed—namely, inflation of zero responses. Second, the participants are observed to switch between the inactive (ZI) and active (non-ZI) regimes throughout the task. Third, the extent of sparseness (or alternatively, activation) in the two processes varies with age and across individuals. EP did get “activate” more frequently at 48 months than at 36 months for some children (e.g., the child depicted in the top panels of Fig. 1) as would be expected developmentally. Instances of PR activation also increased with age for these participants. From a theoretical standpoint, the essence of self-regulation resides in the dependencies between EP and PR as they transition between the inactive and the active regimes. Of particular interest to us are age differences in the way that EP triggers and/or modulates PR, and vice versa.

In this study, OU model was chosen as the starting point because it has been utilized in other contexts to represent self-regulation in adults (Oravec, Tuerlinckx, & Vandekerckhove, 2016). The processes of negative behaviors and emotions and the EP showed random entangling fluctuations around their equilibria. However, due to rare occurrences of some of the EP and PR markers (as related, e.g., to the emerging nature of many young children’s EP), subsequent aggregation of these binary codes yields high proportions of zero responses in the bivariate time-series data. The estimation of the unique equilibrium of the OU process is likely to be biased by the large amount of inactive time points, resulting in an estimated equilibrium that does not reflect the central location of either the inactive or the active regime. In addition, the OU model is not designed to accommodate transitions between two regimes that are characterized by very distinct equilibria, nor does it capture the dependencies between how EP and PR processes transition between these hypothesized regimes—all questions that are of direct interest to the study of self-regulation development in children. These data characteristics and limitations of the OU model motivated our development of a dynamic model for self-regulation in the presence of ZI.

2. Zero-Inflated Ornstein–Uhlenbeck (ZI-OU) Model

We first introduce the OU model. Then, the distinct regime indicators for all processes are incorporated into a bivariate OU model to allow the coefficients of each process to change according to distinct regimes, resulting in the ZI-OU model. Then, we describe the latent regime transition model that governs the dynamics of the regime indicators and allows for subject-specific difference among individuals.

2.1. The OU Model

The OU process is widely used to model stochastic processes that fluctuate around an equilibrium. The SDE representation of OU process is

$$\frac{dx_i(t)}{dt} = \beta(\mu - x_i(t)) + \sigma dw_i(t), \quad (1)$$

where i indexes child in our application and t indexes continuous time. μ represents the equilibrium; σ is a diffusion parameter that quantifies the amount of random fluctuations; and $\beta \geq 0$ is the approach rate toward the equilibrium. Larger β indicates the process approaches to the

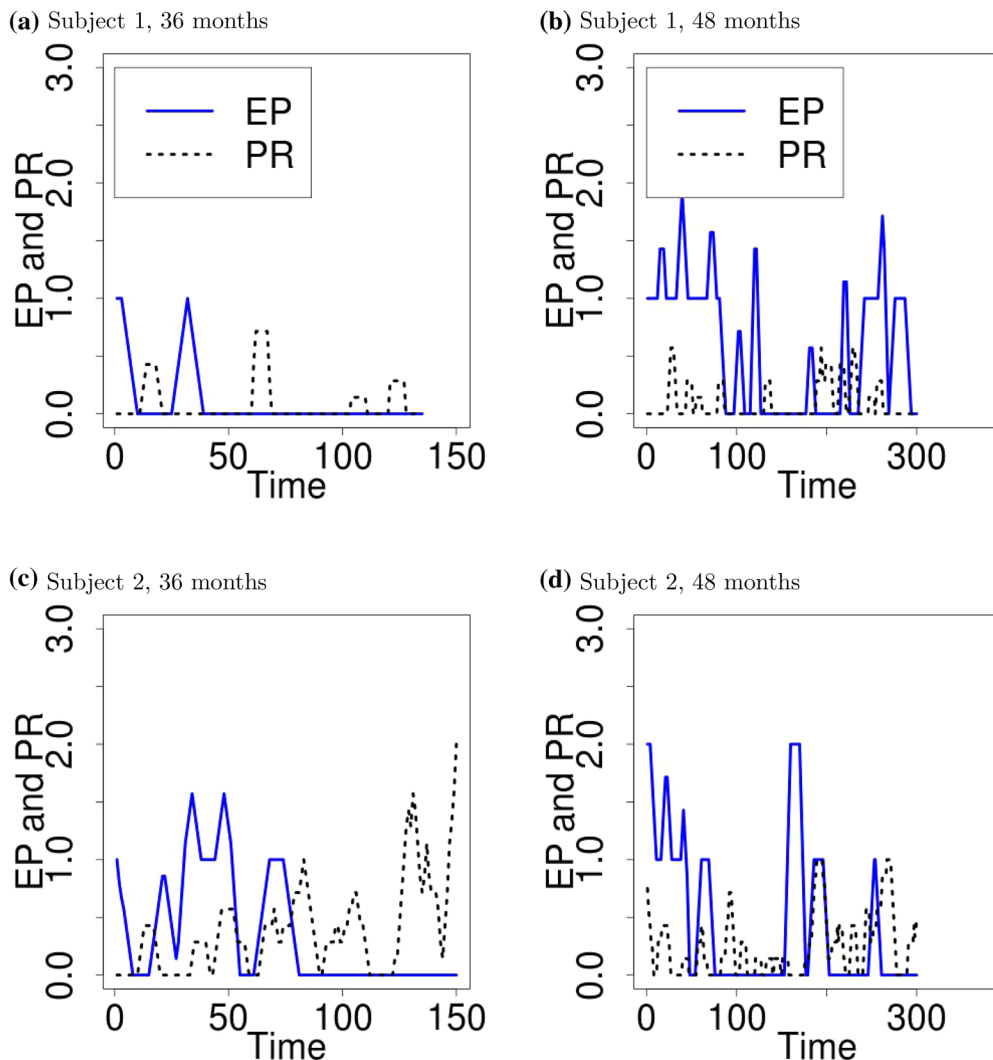


FIGURE 1.

Observed data trajectories of two randomly selected participants at 36 and 48 months. Solid and dashed curves represent EP and PR, respectively.

equilibrium faster. $w_i(t)$ is a standard Wiener process, and the increment, $dw_i(t)$, follows a Gaussian distribution with zero mean and variance that is proportional to the length of time interval, dt . Figure 2a, d shows four different simulated realizations of the OU process with four distinct combinations of parameter values. It can be seen that as β increases (becomes more positive), the process approaches in equilibrium at μ more quickly. With larger σ , a greater range of data values are observed. The rate of change of the OU process is a combination of the deterministic drift function and the random diffusion.

The OU model was utilized in other contexts to represent adults self-regulation where each process fluctuates around an equilibrium (Oravecz et al., 2016). However, the OU model is not likely to adequately characterize the child self-regulation process in the presence of considerable

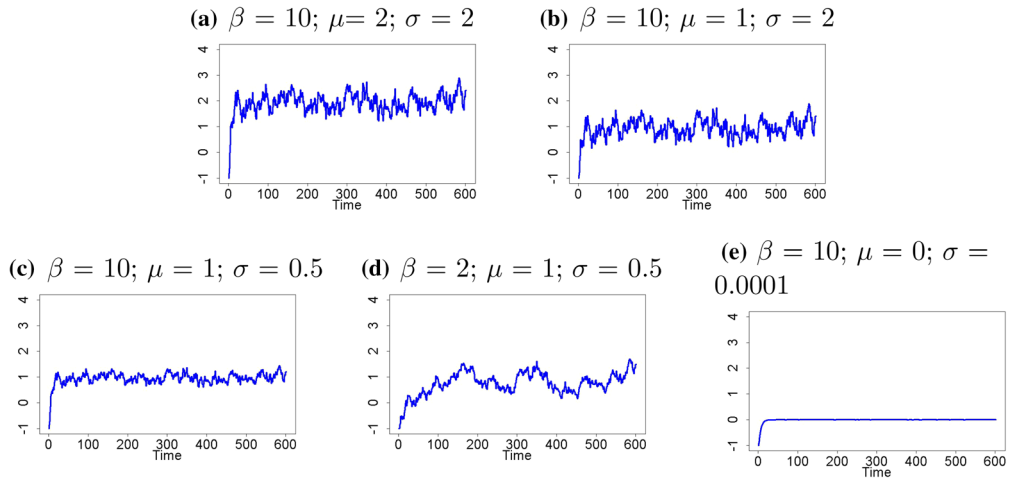


FIGURE 2.

Simulated trajectories of OU processes given various parameter values. β —approaching parameter; μ —equilibrium; σ —diffusion.

time points of inactive state in Fig. 1, which motivated us to propose the ZI extension to the classical OU model.

2.2. OU Model under Extended Non-occurrence of Behavior

The key behind our proposed ZI-OU model is that while a particular process of interest is in a “inactive state,” the corresponding trajectories with repeated occurrences of zeros can essentially be obtained as a special case of the classical OU model in (1), in particular, when $\mu = 0$ and σ is close to 0. A simulated trajectory given the OU model and these parameters is shown in Fig. 2e. In contrast, a child’s affective dynamics while in an “active” state have been modeled, as well, using the classical OU model, but typically with β and σ both greater than zero. In stark contrast to standard mixture models, transition between the active and inactive state occurs within individuals even though there are inter-individual differences in the points at each the transitions occur.

2.3. Regime-Dependent OU Model

To model the dependencies between EP and PR in our sample of young children in the presence of within-individual manifestations of ZI, we propose to incorporate regime switching into the OU model, which allows the OU process to be governed by different sets of dynamic parameters over time and leading to changing process dynamics. One dynamic regime is restricted to show “inactive” dynamics. Specifically, we added a latent regime indicator, $l_i^j(t)$, to each of $j = \text{EP}$ and PR that indicates whether individual i ’s process j is in the active or inactive regime. Two latent regime indicators, $l_i^{\text{EP}}(t)$ and $l_i^{\text{PR}}(t)$, are incorporated into a bivariate OU model to mark the respective regime associated with the EP and PR processes, respectively, at each time point and for each child. This regime-dependent ZI-OU model is expressed as:

$$\frac{d\text{EP}_i(t)}{dt} = \beta_{\text{EP}, l_i^{\text{EP}}(t)} \left(\mu_{\text{EP}, l_i^{\text{EP}}(t)} - \text{EP}_i(t) \right) + \sigma_{\text{EP}, l_i^{\text{EP}}(t)} dw_{\text{EP}, i}(t), \quad (2)$$

$$\frac{d\text{PR}_i(t)}{dt} = \beta_{\text{PR}, l_i^{\text{PR}}(t)} \left(\mu_{\text{PR}, l_i^{\text{PR}}(t)} - \text{PR}_i(t) \right) + \sigma_{\text{PR}, l_i^{\text{PR}}(t)} dw_{\text{PR}, i}(t), \quad (3)$$

where j indexes the two processes EP and PR ($j = \text{EP or PR}$) as explained earlier. $l_i^j(t)$ is the latent regime governing individual i 's process j at time t ; $\mu_{j,l_i^j(t)}$ represents the equilibrium of the j th process for person i at time t ; $\beta_{j,l_i^j(t)}$ is the approach rate of the j th process for person i at time t toward the equilibrium; $\sigma_{j,l_i^j(t)}$ is a diffusion parameter that quantifies the amount of random fluctuations in process j at time t . The parameters $\mu_{j,l_i^j(t)}$, $\beta_{j,l_i^j(t)}$ and $\sigma_{j,l_i^j(t)}$ vary over participants and time only as contingent on $l_i^j(t)$. $w_{\text{EP},i}(t)$ and $dw_{\text{PR},i}(t)$ are standard Wiener processes and the increments. $dw_{\text{EP},i}(t)$ and $dw_{\text{PR},i}(t)$ follow Gaussian distribution with zero means and variances that are proportional to the length of time interval, dt .

Compared to traditional OU process, the regime indicator, $l_i^j(t)$, allows EP and PR to each switch between the inactive and activate regimes to accommodate the inflation of zero responses in the empirical measurements. We define the first regime, $l_i^j(t) = 1$, to be the inactive (ZI) regime for $j = \text{EP, PR}$, respectively, by specifying $\mu_{j,1} = 0$ and $\sigma_{j,1}$ to be a small constant close to 0, such as 0.01. $\beta_{j,1}$ and the parameters for the active regime, i.e., $l_i^j(t) = 2$, are freely estimated. The two regime values for EP and PR mean that the bivariate process can assume one of four possible combinations of regime values: [EP active, PR active]; [EP active, PR inactive]; [EP inactive, PR active]; and [EP inactive, PR inactive].

2.4. Latent Regime Transition Model

The model in Eqs. (2) and (3) only specify the respective dynamics of the system as a whole while in the active and inactive regimes. It does not specify the time evolution of the latent regime indicators or in other words, how a child transitions from an active to inactive regime and vice versa. To represent the switching between the inactive and active regimes of EP and PR, the latent regime indicators are assumed to follow a logistic model that posits that the future regime for each of the two modeled processes (EP and PR) depends both on the current regime of EP and on the current regime of PR as:

$$\begin{aligned}
 p\left(l_{t_{i,k+1}}^{\text{EP}} = r_1 | l_{t_{i,k}}^{\text{EP}} = s_1, l_{t_{i,k}}^{\text{PR}} = s_2\right) &= \frac{\exp\left\{\alpha_{r_1 s_1 s_2}^{\text{EP}T} \mathbf{u}_i\right\}}{\sum_{r=1}^2 \exp\left\{\alpha_{r s_1 s_2}^{\text{EP}T} \mathbf{u}_i\right\}}, \\
 p\left(l_{t_{i,k+1}}^{\text{PR}} = r_2 | l_{t_{i,k}}^{\text{EP}} = s_1, l_{t_{i,k}}^{\text{PR}} = s_2\right) &= \frac{\exp\left\{\alpha_{r_2 s_1 s_2}^{\text{PR}T} \mathbf{u}_i\right\}}{\sum_{r=1}^2 \exp\left\{\alpha_{r s_1 s_2}^{\text{PR}T} \mathbf{u}_i\right\}}, \tag{4}
 \end{aligned}$$

where r_1, r_2, s_1 , and s_2 are indexes for the regimes of EP and PR at the next time point and the current regimes of EP and PR, respectively; they may assume the value of 1 or 2, corresponding to the inactive and active regimes, respectively. The subject-specific covariates in \mathbf{u}_i are used to explain inter-individual-level differences in the probability of being in a particular regime, and regime transition patterns therein. In our empirical illustration, \mathbf{u}_i consists of a constant of 1 and the age of child i at the time of assessment. Thus, $\mathbf{u}_i = (1, \text{age}_i)^T$, in which age_i is a dummy-coded covariate with values of 0 for children at 36 months and 1 for children at 48 months; $\alpha_{r_1 s_1 s_2}^{\text{EP}}$ and $\alpha_{r_2 s_1 s_2}^{\text{PR}}$ are the corresponding vectors of coefficients for these covariates in predicting $l_{t_{i,k+1}}^{\text{EP}} = r_1$ and $l_{t_{i,k+1}}^{\text{PR}} = r_2$ given that $l_{t_{i,k}}^{\text{EP}} = s_1$ and $l_{t_{i,k}}^{\text{PR}} = s_2$. Thus, we expect the probability for the i th child's j th process to transition into regime r_j at time $t_{i,k+1}$ to depend on the child's age at time $t_{i,k}$, its regime at the current observed time point (s_j), and the regime of the opposing process at the current observed time point ($s_{j'}, j' \neq j$).

In sum, the regime transition model in Eq. (4) dictates that for each of EP and PR, there is a 4×2 transition matrix that describes the log-odds (and by extension, probability) of being in an active (vs. inactive) regime next conditional on its own current regime and the opposing process's current

TABLE 1.
Log-odds parameters under different statuses of EP and PR at the current and previous time points.

		Log-odds parameters		Full main effects		Constrained main effects	
EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	EP _{<i>t_{i,k+1}</i>}	EP _{<i>t_{i,k+1}</i>}	EP _{<i>t_{i,k+1}</i>}	EP _{<i>t_{i,k+1}</i>}	EP _{<i>t_{i,k+1}</i>}	EP _{<i>t_{i,k+1}</i>}
		Inactive	Active	Inactive	Active	Inactive	Active
Inactive	Inactive	$\alpha_{111,0}^{EP}$	$\alpha_{211,0}^{EP}$	$\alpha_{111,0}^{EP}$	0	$\alpha_{111,0}^{EP}$	0
Inactive	Active	$\alpha_{112,0}^{EP}$	$\alpha_{212,0}^{EP}$	$\alpha_{112,0}^{EP}$	0	$\alpha_{112,0}^{EP}$	0
Active	Inactive	$\alpha_{121,0}^{EP}$	$\alpha_{221,0}^{EP}$	0	$\alpha_{221,0}^{EP}$	0	$\alpha_{221,0}^{EP}$
Active	Active	$\alpha_{122,0}^{EP}$	$\alpha_{222,0}^{EP}$	0	$\alpha_{222,0}^{EP}$	0	$\alpha_{222,0}^{EP}$

EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k+1}</i>}	PR _{<i>t_{i,k+1}</i>}	PR _{<i>t_{i,k+1}</i>}	PR _{<i>t_{i,k+1}</i>}	PR _{<i>t_{i,k+1}</i>}	PR _{<i>t_{i,k+1}</i>}
		Inactive	Active	Inactive	Active	Inactive	Active
Inactive	Inactive	$\alpha_{111,0}^{PR}$	$\alpha_{211,0}^{PR}$	$\alpha_{111,0}^{PR}$	0	$\alpha_{111,0}^{PR}$	0
Inactive	Active	$\alpha_{112,0}^{PR}$	$\alpha_{212,0}^{PR a}$	0	$\alpha_{212,0}^{PR a}$	0	$\alpha_{2,2,0}^{PR}$
Active	Inactive	$\alpha_{121,0}^{PR}$	$\alpha_{221,0}^{PR}$	$\alpha_{121,0}^{PR}$	0	$\alpha_{121,0}^{PR}$	0
Active	Active	$\alpha_{122,0}^{PR}$	$\alpha_{222,0}^{PR a}$	0	$\alpha_{222,0}^{PR a}$	0	$\alpha_{2,2,0}^{PR}$

The zeros are parameters fixed for identification.
^a $\alpha_{212,0}^{PR}$ and $\alpha_{222,0}^{PR}$ are set to be identical as $\alpha_{2,2,0}^{PR}$.

regime. This leads to a transition log-odds matrix and, correspondingly, a transition probability matrix of size 8×2 for both EP and PR, as shown in Table 1 under column “log-Odds parameters.” To identify the regime transition model, at least one cell in each of the 8 rows of this transition matrix has to be set to a constant—typically zero—to identify the model. One plausible option is to set the coefficients in cells that suggest a change of regime at the next time point (e.g., switching from active to inactive) to zero regardless of the current regime of the opposing process. These cells are the cells containing 0 in Table 1 under the column, “full main effects.” That is, for these cells, the corresponding log-odd parameters in α_{121}^{EP} , α_{211}^{EP} , α_{122}^{EP} , α_{212}^{EP} , α_{112}^{PR} , α_{211}^{PR} , α_{122}^{PR} , and α_{221}^{PR} are all set to 0, a vector of zeros of appropriate dimension. As an example, in the model with age effects, the subscripts in $(\cdot)_{121}$ correspond to EP being in the inactive regime at the next time point ($r_1 = 1$) given that EP and PR are active and inactive, respectively, at the current time point (i.e., $s_1 = 2$ and $s_2 = 1$). Thus, $p(l_{t_i,k+1}^{EP} = 1 | l_{t_i,k}^{EP} = 2, l_{t_i,k}^{PR} = 1)$ is the equivalent of the first entry in the third row under “full main effects” in Table 1, shown as a zero. This is because the corresponding coefficients that predict this particular log-odd, α_{121}^{EP} , including the intercept, $\alpha_{121,0}^{EP}$, and a regression coefficient associated with age, $\alpha_{121,1}^{EP}$, are both set to zeros for identification purposes.

The transitional probabilities can then be calculated as log-odds (LO). For example, under the “reduced main and age effects” model in Table 2,

$$p\left(l_{t_i,k+1}^{EP} = 1 | l_{t_i,k}^{EP} = 1, l_{t_i,k}^{PR} = 1, \text{age}_i = 36m\right) = \frac{\exp\left(\alpha_{111,0}^{EP}\right)}{\exp\left(\alpha_{111,0}^{EP}\right) + \exp(0)},$$

$$p\left(l_{t_i,k+1}^{EP} = 2 | l_{t_i,k}^{EP} = 1, l_{t_i,k}^{PR} = 1, \text{age}_i = 36m\right) = \frac{\exp(0)}{\exp\left(\alpha_{111,0}^{EP}\right) + \exp(0)},$$

TABLE 2.
Log-odds parameters under different statuses of EP and PR at the current and previous time points.

Full age effects		36 months		48 months		Equations
EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	EP _{<i>t_{i,k+1}</i>} Inactive	EP _{<i>t_{i,k+1}</i>} Active	EP _{<i>t_{i,k+1}</i>} Inactive	EP _{<i>t_{i,k+1}</i>} Active	
Inactive	Inactive	$\alpha_{111,0}^{EP}$	0	$\alpha_{111,0}^{EP} + \alpha_{111,1}^{EP}$	0	
Inactive	Active	$\alpha_{112,0}^{EP}$	0	$\alpha_{112,0}^{EP} + \alpha_{112,1}^{EP}$	0	
Active	Inactive	0	$\alpha_{221,0}^{EP}$	0	$\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP}$	
Active	Active	0	$\alpha_{222,0}^{EP}$	0	$\alpha_{222,0}^{EP} + \alpha_{222,1}^{EP}$	

EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k+1}</i>} Inactive	PR _{<i>t_{i,k+1}</i>} Active	PR _{<i>t_{i,k+1}</i>} Inactive	PR _{<i>t_{i,k+1}</i>} Active
Inactive	Inactive	$\alpha_{111,0}^{PR}$	0	$\alpha_{111,0}^{PR} + \alpha_{111,1}^{PR}$	0
Inactive	Active	0	$\alpha_{2,2,0}^{PR}$	0	$\alpha_{2,2,0}^{PR} + \alpha_{212,1}^{PR}$
Active	Inactive	$\alpha_{121,0}^{PR}$	0	$\alpha_{121,0}^{PR} + \alpha_{121,1}^{PR}$	0
Active	Active	0	$\alpha_{2,2,0}^{PR}$	0	$\alpha_{2,2,0}^{PR} + \alpha_{222,1}^{PR}$

Reduced main and age effects		36 months		48 months		Equations
EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	EP _{<i>t_{i,k+1}</i>} Inactive	EP _{<i>t_{i,k+1}</i>} Active	EP _{<i>t_{i,k+1}</i>} Inactive	EP _{<i>t_{i,k+1}</i>} Active	
Inactive	Inactive	$\alpha_{111,0}^{EP}$	0	$\alpha_{111,0}^{EP}$	0	(8)
Inactive	Active	$\alpha_{112,0}^{EP}$	0	$\alpha_{112,0}^{EP}$	0	(9)
Active	Inactive	0	$\alpha_{221,0}^{EP}$	0	$\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP}$	(10)
Active	Active	0	$\alpha_{222,0}^{EP}$	0	$\alpha_{222,0}^{EP}$	(11)

EP _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k}</i>}	PR _{<i>t_{i,k+1}</i>} Inactive	PR _{<i>t_{i,k+1}</i>} Active	PR _{<i>t_{i,k+1}</i>} Inactive	PR _{<i>t_{i,k+1}</i>} Active	Equations
Inactive	Inactive	$\alpha_{1,1,0}^{PR}$	0	$\alpha_{1,1,0}^{PR} + \alpha_{111,1}^{PR}$	0	(12)
Inactive	Active	0	$\alpha_{2,2,0}^{PR}$	0	$\alpha_{2,2,0}^{PR}$	(14)
Active	Inactive	$\alpha_{1,1,0}^{PR}$	0	$\alpha_{1,1,0}^{PR} + \alpha_{121,1}^{PR}$	0	(13)
Active	Active	0	$\alpha_{2,2,0}^{PR}$	0	$\alpha_{2,2,0}^{PR}$	(15)

The zeros are parameters fixed for identification. The constrained parameters are those with “.” in the subscript.

$$p \left(l_{t_i,k+1}^{EP} = 1 | l_{t_i,k}^{EP} = 2, l_{t_i,k}^{PR} = 1, \text{age}_i = 48m \right) = \frac{\exp(0)}{\exp(0) + \exp \left(\alpha_{221,0}^{PR} + \alpha_{221,1}^{PR} \right)}, \text{ and}$$

$$p \left(l_{t_i,k+1}^{EP} = 2 | l_{t_i,k}^{EP} = 2, l_{t_i,k}^{PR} = 1, \text{age}_i = 48m \right) = \frac{\exp \left(\alpha_{221,0}^{PR} + \alpha_{221,1}^{PR} \right)}{\exp(0) + \exp \left(\alpha_{221,0}^{PR} + \alpha_{221,1}^{PR} \right)}.$$

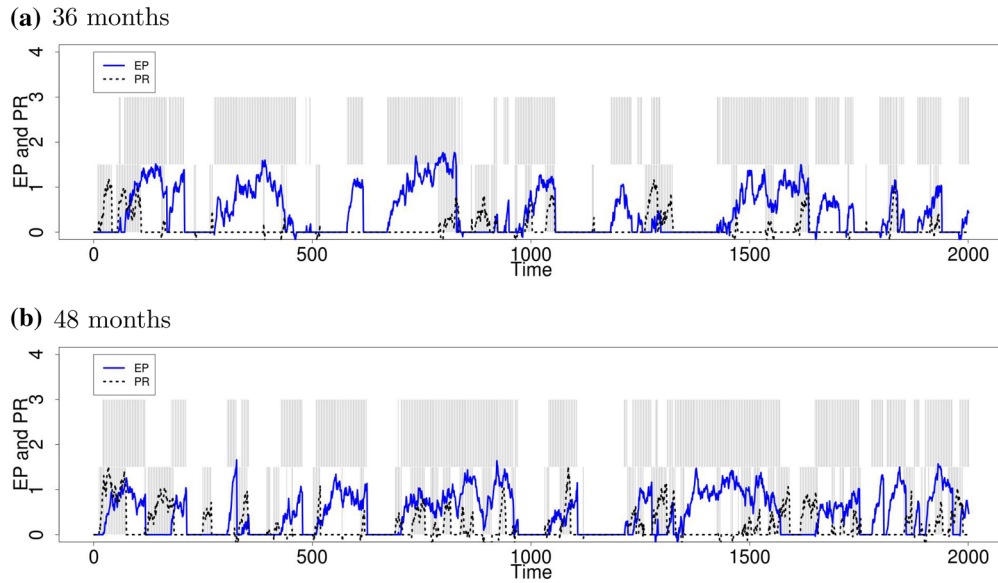


FIGURE 3.

Simulated trajectories of EP (solid) and PR (dashed) based on the ZI-OU at 36 months (top) and 48 months (bottom), respectively. The stacked shaded regions mark portions of the data during which EP (upper shaded region) and PR (lower shaded region) are active, respectively.

Simulated trajectories based on this example and the parameters estimated in the empirical data are shown in Fig. 3. The simulated trajectories demonstrate that the proposed ZI-OU model may be a plausible model for the empirical data plotted in Fig. 1 in that it helps capture the following aspects of the data. First, the two hypothesized processes are allowed to switch between inactive (zero responses) and active phases recurrently at individual-specific and time-varying intervals. Second, the status of one process influences the status of the other process. Third, some age differences can be observed in the dependencies between the EP and PR processes over time. In summary, Eqs. (2)–(4) collectively constitute our entire proposed ZI-OU model.

3. Regime-Switching Stochastic Differential Model

The proposed ZI-OU model is a special case of a more general regime-switching SDE model as

$$d\mathbf{x}_i(t) = \mathbf{f}(\mathbf{x}_i(t), \boldsymbol{\theta}_{\mathbf{I}_i(t)})dt + \mathbf{S}(\mathbf{x}_i(t), \boldsymbol{\theta}_{\mathbf{I}_i(t)})d\mathbf{w}_i(t), \quad (5)$$

where $\mathbf{x}_i(t)$ is a vector of latent process variables of interest (e.g., a child's latent EP and PR), $\mathbf{f}(\cdot) = (f_1(\cdot), \dots, f_q(\cdot))$ is a $q \times 1$ vector of drift functions, \mathbf{S} is a $q \times q$ matrix of diffusion functions, and $\mathbf{w}_i(t)$ is a $q \times 1$ vector of standard Wiener processes, whose differentials, $d\mathbf{w}_i(t)$, are Gaussian distributed with zero means and variances that increase with the length of time interval, dt . $\boldsymbol{\theta}_{\mathbf{I}_i(t)}$ is a vector of parameters governing the dynamics of $\mathbf{x}_i(t)$, and whose values depend on the latent regime characterizing person i at time t . Let $\mathbf{I}_i(t) = (I_{ji}(t); j = 1, \dots, q)$ denote a vector of latent regime indicators showing the regime of the j th element in \mathbf{x} for the i th child at time t for a total of R regimes. In our empirical application, $q = 2$, corresponding to the two latent variables in \mathbf{x} : EP and PR; $R = 2$, corresponding to a total of two hypothesized regimes (inactive and active regimes).

To model the patterns of transition among different regimes, we generalized Eq. (4) and assume a Markovian transition model such that the probability for the i th child's j th process to be in regime r at time $t_{i,k}$ depends on person-specific covariates and the child's previous regime for processes $1, \dots, q$ as:

$$P(l_{jt_{i,k+1}} = r | \mathbf{l}_{t_{i,k}} = \mathbf{s}) = p_{ijr,\mathbf{s}} = \frac{\exp\{\boldsymbol{\alpha}_{rs}^{(j)T} \mathbf{u}_i\}}{\sum_{r=1}^R \exp\{\boldsymbol{\alpha}_{rs}^{(j)T} \mathbf{u}_i\}}, \quad (6)$$

where $r = 1, \dots, R$ is the index of regime at the next time point for the j th process, $\mathbf{s} = (s_1, \dots, s_q)$ is a vector containing the “lag-one” indicator of the child's regime membership for all q processes at the current time point, and $p_{ijr,\mathbf{s}}$ is the transition probability satisfying $\sum_r^R p_{ijr,\mathbf{s}} = 1$ for all \mathbf{s} . The transition probability $p_{ijr,\mathbf{s}}$ is subject-specific as it is a function of \mathbf{u}_i , a $m \times 1$ vector of person-specific covariates; and $\boldsymbol{\alpha}_{rs}^{(j)}$, the associated vector of regression coefficients. The term $\{\boldsymbol{\alpha}_{rs}^{(j)T} \mathbf{u}_i\}$ in Eq. (6) denotes the LO for the j th process to transition into regime r at time $t_{i,k+1}$ given the state at the current time point is \mathbf{s} . In our empirical application $\mathbf{s} = (s_1, s_2)$.

Model (6) allows the current regime for a particular latent process to depend not only on the previous regime of that same process, but also the previous regime of other related processes. For instance, in our motivating example, the current regime of EP may depend both on EP's previous regime and on the previous regime of the opposing process, PR. The strength of such dependency is governed by the parameters in $\boldsymbol{\alpha}_{rs}^{(j)}$. For instance, the probability for child i 's EP to be in regime r at next time $t_{i,k+1}$ is allowed to differ depending on whether the child's EP was currently active (as indicated by s_1 , a binary indicator of the child's current regime for EP at $t_{i,k}$), as well as whether the child's PR was current active (as indicated by s_2 , a binary indicator of the child's current regime for PR at $t_{i,k}$).

Some regime-switching parameters in $\boldsymbol{\alpha}_{rs}^{(j)}$ shown in Eq. (6) need to be fixed for identification purposes similar to those in Eq. (4) for the motivating example. Specifically, for the j th process, conditional on $\mathbf{s} = (s_1, \dots, s_q)$, the probability of switching to one of the R regimes has to be fixed by setting all the parameters in $\boldsymbol{\alpha}_{rs}$ linked to that regime—which serves now as the “reference” regime—to zero.

The proposed regime-switching SDE framework generalizes the conventional SDE framework by allowing the parameters in an SDE to transition among different values depending on the operating regime at a particular time point. In addition, the general model in Eqs. (5) and (6) includes a different regime indicator for each element in the dynamic process \mathbf{x} , thus facilitating the analysis of interdependent multivariate dynamic processes and their associated transition patterns.

4. Estimation and Inference

4.1. Numerical Solution of SDE

In most empirical studies, we only measure the dynamic processes at time points $t_{i,k}$ for $k = 1, \dots, T_i$ and $i = 1, \dots, n$, where $t_{i,k}$ is the k th time point for the i th individual. Also, most SDEs do not have analytical solutions, but rather have to be approximated by numerical solutions for estimation and inferential purposes. We used Euler–Maruyama approximation (Kloeden & Platen, 1999), which provides a discretized approximation at selected time points. For the model in (2)–(3), a first-order Euler approximation to the solution may be obtained as

$$\begin{aligned} \text{EP}_{i,t_i,k+1} &= \text{EP}_{i,t_i,k} + \beta_{\text{EP},l_{i,k}^{\text{EP}}} (\mu_{\text{EP},l_{i,k}^{\text{EP}}} - \text{EP}_{i,t_i,k}) \Delta t_{i,k} + \Delta t_{i,k}^{1/2} \sigma_{\text{EP},l_{i,k}^{\text{EP}}} z_{\text{EP},i,t_i,k}, \\ \text{PR}_{i,t_i,k+1} &= \text{PR}_{i,t_i,k} + \beta_{\text{PR},l_{i,k}^{\text{PR}}} (\mu_{\text{PR},l_{i,k}^{\text{PR}}} - \text{PR}_{i,t_i,k}) \Delta t_{i,k} + \Delta t_{i,k}^{1/2} \sigma_{\text{PR},l_{i,k}^{\text{PR}}} z_{\text{PR},i,t_i,k}, \end{aligned}$$

where $\Delta t_{i,k} = t_{i,k+1} - t_{i,k}$, $z_{\text{EP},i,t_i,k}$ and $z_{\text{PR},i,t_i,k}$ are independent standard normal random variables, and $l_{i,k}^{\text{EP}}$ and $l_{i,k}^{\text{PR}}$ are the latent regime indicator of EP and PR at time $t_{i,k}$ for the i th subject, respectively. It is worth noting that l also depends on the subject index i in addition to $t_{i,k}$. We omit i for notational simplicity.

The discretized approximation of the SDEs in (5) at selected time points $t_{i,k}$ for $k = 1, \dots, T_i$ and $i = 1, \dots, n$ is

$$\Delta \mathbf{x}_{i,t_i,k} = \mathbf{f}(\mathbf{x}_{i,t_i,k}, \boldsymbol{\theta}_{i,t_i,k}) \Delta t_{i,k} + \Delta t_{i,k}^{1/2} \mathbf{S}(\mathbf{x}_{i,t_i,k}, \boldsymbol{\theta}_{i,t_i,k}) \mathbf{z}_{i,t_i,k},$$

where $\Delta \mathbf{x}_{i,t_i,k} = \mathbf{x}_{i,t_i,k+1} - \mathbf{x}_{i,t_i,k}$, $\mathbf{l}_{i,k} = (l_{1,t_i,k}, \dots, l_{q,t_i,k})$ is the processes and latent regime indicators at the time which the k th observation of the i th subject was observed, $\Delta t_{i,k} = t_{i,k+1} - t_{i,k}$ for $0 \leq k < T_i$ and $1 \leq i \leq n$; $\mathbf{z}_{i,t_i,k}$ conforms to a multivariate Gaussian distribution, $N(\mathbf{0}, \mathbf{I}_q)$, in which \mathbf{I}_q is a $q \times q$ identity matrix. When $k = 0$, the initial conditions of the processes in \mathbf{x} , denoted as \mathbf{x}_{i0} , are assumed to be known for all i .

4.2. Bayesian Estimation and Inference

Bayesian approaches are useful tools for the estimation and inference for SDEs (Durham & Gallant, 2002; Elerian, Chib, & Shephard, 2001; Golightly & Wilkinson, 2008; Roberts & Stramer, 2001). Bayesian methods have also been applied to fit regime-switching models (Ghysels, McCulloch, & Tsay, 1998; Kim & Kim, 2015; Yümlü, Gürgen, Cemgil, & Okay, 2015). The Bayesian methods may provide more modeling flexibility (Calvet & Fisher, 2004; Fox, Sudderth, Jordan, & Willsky, 2010) and information to explore multiple local maxima. Markov chain Monte Carlo (MCMC) algorithms (Geman & Geman, 1984; Hastings, 1970) are used to generate samples from the joint posterior distribution of all parameters and latent variables, which are accomplished through the JAGS software (Plummer, 2003). Sample JAGS code for implementing the model in the motivating example is included in ‘‘Appendix A.’’

As distinct from frequentist approaches wherein the parameters in a model are treated as fixed and unknown, Bayesian approaches consider the parameters as random variables. Their distributions are quantified by prior distributions before any information from the observed data is incorporated. The combination of information of the prior distribution and the data likelihood leads to the posterior distributions, based on which the estimation and inference are obtained. We used the following prior distributions for the parameters in (2)–(3) for $l = 2$

$$\begin{aligned} \beta_{\text{EP},l} &\sim N(\beta_{\text{EP},0l}, \sigma_{\beta_{\text{EP},l}}^2) I(0, \infty), & \beta_{\text{PR},l} &\sim N(\beta_{\text{PR},0l}, \sigma_{\beta_{\text{PR},l}}^2) I(0, \infty), \\ \mu_{\text{EP},l} &\sim N(\mu_{\text{EP},0l}, \sigma_{\mu_{\text{EP},l}}^2), & \mu_{\text{PR},l} &\sim N(\mu_{\text{PR},0l}, \sigma_{\mu_{\text{PR},l}}^2), \\ \sigma_{\text{EP},l} &\sim IG(a_{\text{EP},1l}, a_{\text{EP},2l}), & \sigma_{\text{PR},l} &\sim IG(a_{\text{PR},1l}, a_{\text{PR},2l}) \end{aligned} \quad (7)$$

and these prior distributions for the parameters in (4)

$$\boldsymbol{\alpha}_{r_1 s_1 s_2}^{\text{EP}} \sim N(\boldsymbol{\alpha}_{0r_1 s_1 s_2}^{\text{EP}}, \boldsymbol{\Sigma}_{\text{EP},0r_1 s_1 s_2}), \boldsymbol{\alpha}_{r_2 s_1 s_2}^{\text{PR}} \sim N(\boldsymbol{\alpha}_{0r_2 s_1 s_2}^{\text{PR}}, \boldsymbol{\Sigma}_{\text{PR},0r_2 s_1 s_2}), \quad (8)$$

where $\beta_{j,0l}$, $\sigma_{\beta_{j,l}}^2$, $\mu_{j,0l}$, $\sigma_{\mu_{j,l}}^2$, $a_{j,1l}$, $a_{j,2l}$, $\alpha_{j,0r_1s_1s_2}$, and positive definite matrix $\Sigma_{j,0r_1s_1s_2}$, for $j = \text{EP, PR}$ are hyperparameters, the values of which are assumed to be given by prior information. We note again that r_1, s_1, s_2 may be 1 or 2. The prior distributions (8) only apply to the parameters that are not fixed for identification. $N(\cdot, \cdot)$, $N(\cdot, \cdot)I(\cdot, \cdot)$, and $IG(\cdot, \cdot)$ are normal distribution, truncated normal distribution, and inverse gamma distribution, respectively.

These prior distributions are conjugate in the sense that the distribution families of the full conditional distributions (“Appendix B”) of the parameters are the same as those of the prior distributions. They were selected mainly for simplicity and computational efficiency because under conjugate priors, the conditional distributions are of known forms and the Gibbs sampler (Geman & Geman, 1984) can be used to ease sampling. When other distributions are used, alternative MCMC algorithms may be employed, some candidates of which include the slice sampler (Neal, 2003) and adaptive rejection Metropolis sampling algorithm (Gilks, Best, & Tan, 1995), but at the cost of increased computational burden.

The prior distributions affect the posterior distribution in a Bayesian setting. However, when the information of the prior distribution is not very strong and the sample size is large enough, the impact on the posterior distribution is usually not substantial. To ensure that this was the case under the model and sample size configuration considered, we evaluated the impact of different hyperparameters of the prior distributions in the simulation study.

To obtain the initial values for the latent regime indicator, we used a simpler mixture model to help cluster EP and PR into two groups for all subject at all time points with the `mclust R` package (Fraley, Raftery, Murphy, & Scrucca, 2012). Missing values of EP and PR were also sampled as latent variables in the MCMC algorithm. We checked the convergence of the MCMC algorithm through the estimated potential scale reduction (EPSR, Gelman, Meng, & Stern, 1996) based on three MCMC chains starting from different initial values for a simulated data set in each condition. After discarding N_0 burn-in samples before the convergence of the MCMC algorithm, the posterior distribution of the parameters in (2)–(4) can be approximated by the empirical distribution of the remaining N_1 samples. In the empirical application and the simulation study, we used $N_0 = 1000$ and $N_1 = 4000$. After the burn-in period, the EPSR values of all parameters were smaller than 1.2, indicating that the three chains have converged. We only used one chain in the complete simulation with massive replicated data sets. The autocorrelations of the MCMC samples were not large, and thinning did not have a significant impact on the parameter estimates. The MCMC sampling with one core of an Intel E5 computer took about nine and five hours for the empirical data analysis and one replication in the simulation study, respectively.

Samples from the empirical approximated posterior distribution were used to perform statistical inference. Sample means and standard deviations of the parameters were used as estimates of the posterior means and standard deviations of the parameters. Many other quantities related to the posterior distributions may be estimated with the MCMC samples. For example, the percentile intervals of the empirical posterior distribution can be used as credible intervals to quantify the uncertainty around the point estimates (Gill, 2014). When the posterior distribution of a parameter is symmetric, we can also compute a pseudo p value with the estimated posterior mean (or median) and standard deviation (the latter is the analogue to standard error in the frequentist framework) of the empirical distribution.

5. Simulation Study

We used the ZI-OU processes shown in Eqs. (2) and (3), and the structure of the regime-switching functions is similar to Eqs. (9)–(16) to perform a targeted simulation to evaluate properties of the Bayesian estimation procedures. As mentioned earlier, some OU-related parameters specific to the inactive (ZI) regime, including $\beta_{\text{EP},1}$ and $\beta_{\text{PR},1}$, did not show satisfactory conver-

TABLE 3.
Bias and standard error based on the estimates of 100 replications given three prior distributions.

	True	Diffuse prior			Informative prior 1			Informative prior 2		
		Bias	SE	Width	Bias	SE	Width	Bias	SE	Width
$\beta_{EP,1}$	2	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001
$\beta_{EP,2}$	0.2	0.000	0.002	0.008	0.000	0.002	0.004	0.000	0.002	0.004
$\mu_{EP,2}$	2	0.001	0.006	0.024	0.000	0.006	0.012	0.001	0.006	0.012
$\sigma_{EP,2}^2$	0.01	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\beta_{PR,1}$	2	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001
$\beta_{PR,2}$	0.2	0.000	0.002	0.007	0.000	0.002	0.004	0.000	0.002	0.004
$\mu_{PR,2}$	2	0.001	0.006	0.022	0.000	0.006	0.011	0.000	0.006	0.011
$\sigma_{PR,2}^2$	0.01	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\alpha_{111,0}^{EP}$	4.2	-0.002	0.078	0.299	-0.018	0.078	0.150	-0.008	0.078	0.149
$\alpha_{121,0}^{EP}$	4.1	0.020	0.075	0.287	0.025	0.075	0.143	0.034	0.075	0.144
$\alpha_{221,0}^{EP}$	3.9	0.024	0.116	0.442	0.019	0.116	0.221	0.037	0.115	0.221
$\alpha_{222,0}^{EP}$	4	-0.001	0.078	0.299	-0.012	0.077	0.149	-0.004	0.078	0.150
$\alpha_{221,1}^{EP}$	0.3	-0.005	0.179	0.684	0.013	0.180	0.344	-0.004	0.178	0.340
$\alpha_{1.1,0}^{PR}$	4.1	0.008	0.079	0.303	0.005	0.078	0.149	0.014	0.079	0.151
$\alpha_{2.2,0}^{PR}$	4.2	0.013	0.058	0.222	0.011	0.058	0.111	0.016	0.058	0.111
$\alpha_{111,1}^{PR}$	-0.3	0.000	0.125	0.479	0.002	0.124	0.235	-0.008	0.125	0.239
$\alpha_{121,1}^{PR}$	-0.3	0.005	0.137	0.526	0.001	0.135	0.258	-0.008	0.136	0.259

True, true values in simulation; Bias, estimates minus true values; SE, standard errors estimated by the standard deviation of MCMC samples from the posterior distribution; Width, the width of the 95% credible intervals.

gence during empirical model fitting and were set to known constants to simplify the model for the empirical data. Here, we freed up these parameters to be estimated to evaluate if they could indeed be uniquely determined. Thus, the parameters in the OU processes were set to $\beta_{EP,1} = \beta_{PR,1} = 2$, $\beta_{EP,2} = \beta_{PR,2} = 0.2$, $\mu_{EP,2} = \mu_{PR,2} = 2$, and $\sigma_{\beta_{EP,2}}^2 = \sigma_{\beta_{PR,2}}^2 = 0.01$. For the inactive component, we fixed $\mu_{EP,1} = \mu_{PR,1} = 0$ and $\sigma_{\beta_{EP,1}}^2 = \sigma_{\beta_{PR,1}}^2 = 0.0001$. The regime-switching functions in the simulation study were set to mirror the final structure from our empirical modeling results, namely (9)–(16), where true values of all log-odds parameters were set according to Table 3. The covariate “age” was simulated from a Bernoulli distribution with a probability parameter of .5.

Diffuse prior distributions were used for the parameters in (7) and (8). Specifically, the means of the prior distributions were set equal to the true values and the prior variance was set to 10^2 for $\beta_{EP,2}$, $\beta_{PR,2}$, $\mu_{EP,2}$, and $\mu_{PR,2}$; $a_{EP,12} = a_{PR,12} = 2$, and $a_{EP,22} = a_{PR,22} = 0.01$. A diffuse univariate normal distribution with mean of 0 and variance 10^4 was assigned to the coefficients in (9)–(16). Using the above simulation setting, we simulated 100 subjects with 400 time points from each subject to mirror the sample size configuration in our motivating example. One hundred

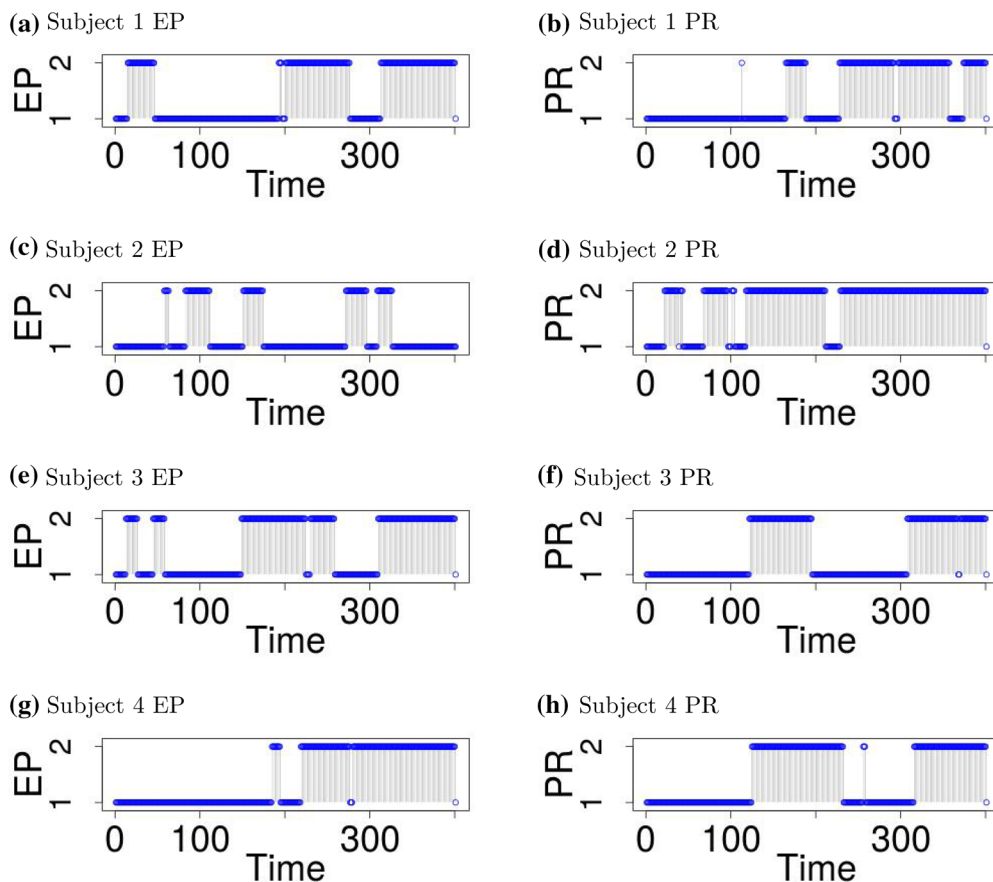


FIGURE 4.

The true and estimated mixture indicators of EP (left) and PR (right) based on the ZI-OU for four subjects in one simulation replication. The shaded region marks portions of the data during which PR and EP are active, respectively. The points represent the estimated regime indicated by the mixture indicators at every time point.

Monte Carlo replications were generated. For each replication, estimation and inference were performed using the default MCMC algorithms implemented in JAGs.

5.1. Simulation Results

The biases and standard error estimates (taken as the standard deviations of the posterior distributions of the parameters) of the parameters across 100 replications are shown in Table 3 under “Diffuse Prior.” We found that all parameters, including those associated with the OU processes in the inactive regime as well as in the active regimes, were all recovered accurately. Moreover, the parameters in regime-switching functions were also estimated with small biases, which provided some evidence that the latent mixture indicators were recovered well for most subjects and time points. Plots of the estimated versus true latent mixture indicators for four randomly selected subjects in one replication, as shown in Fig. 4, provided further verification that the true latent mixture indicators were indeed recovered satisfactorily.

To study the sensitivity of the Bayesian results to the prior choices, we reanalyzed the simulated data with three informative prior settings. To obtain distinct prior distributions, we changed the prior variances to 1 and the prior means to one-half of the true parameter values in Table 3

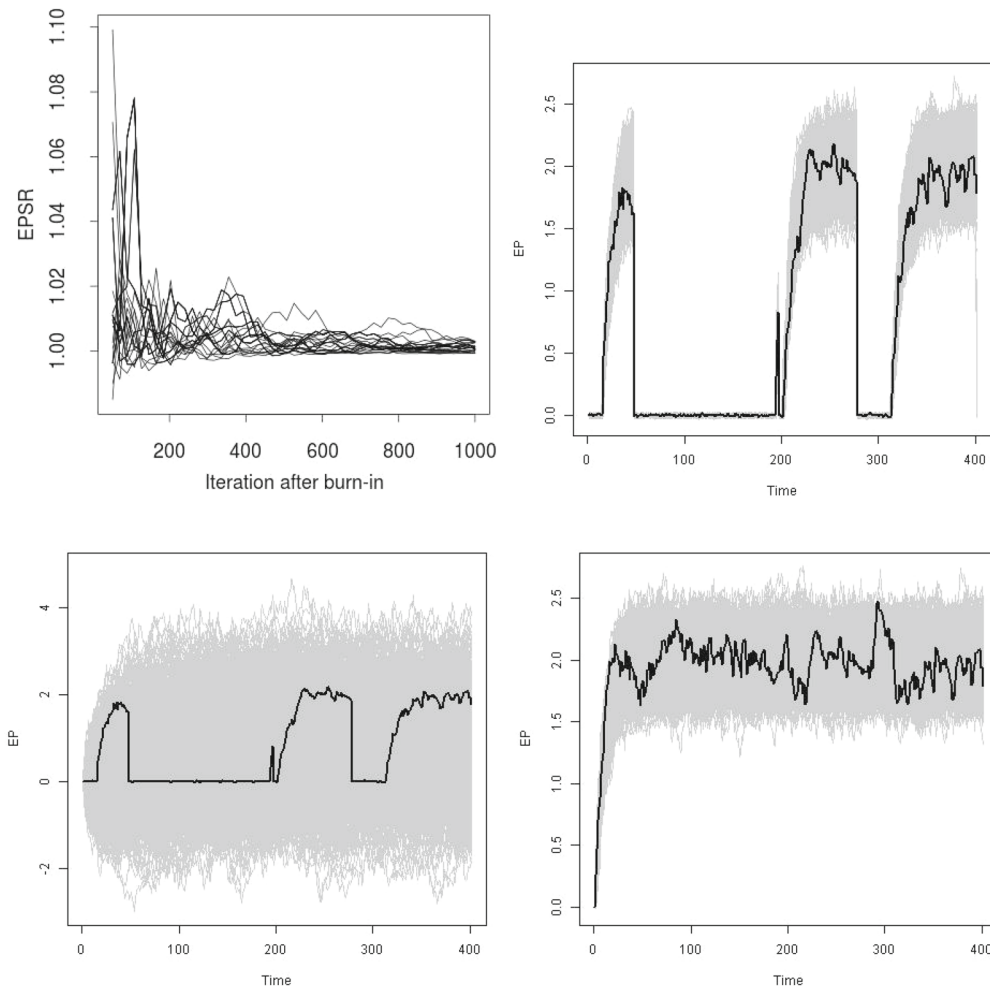


FIGURE 5.

The upper left panel is the EPSR value based on three distinct starting values and a randomly selected simulated data set. The upper right panel shows one randomly selected simulated trajectory (solid) based on the ZI-OU model and the 1000 posterior predictive trajectories (shaded region) based on the estimation of the ZI-OU model. The lower left panel is based on the same data set, and the predictive trajectories are based on the estimation of the classical OU model. The lower right panel is based on data generated from the classical OU model, and the predictive trajectories are based on the estimation of the ZI-OU model.

(Informative Prior 1), the true values and twice the magnitudes of the parameter values (Informative Prior 2), respectively. As the number of subjects and time points are large in the empirical and simulated data sets, the likelihood function dominates the prior distribution and we found that the posterior distribution is not very sensitive to the prior specification. The resultant point and standard error estimates are reported in Table 3, indicating that our modeling results were not very sensitive to our choices of prior specification.

To check the convergence of the MCMC algorithm, we randomly selected one Monte Carlo replication and ran the MCMC algorithms from three distinct sets of starting values for the modeling parameters. The EPSR values (Gelman, 1996) of every parameters based on samples after burn-in and thinned by 4 are shown in the upper left panel of Fig. 5, which indicated that the MCMC chains converged after 1000 iterations.

To check the fit of the proposed ZI-OU model, we generated 1000 posterior predictive trajectories for each subject with the same setting as we generated the simulated data. Each posterior predictive trajectory is based on one MCMC sample of parameter and latent regime indicators. The trajectories of a randomly selected subject (the solid curve) and the corresponding predictive trajectories (shaded regions) are shown in the upper right panel of Fig. 5. The other subjects showed similar patterns. The posterior trajectories showed that the ZI-OU captures the patterns of the regime-switching trajectories accurately.

To demonstrate the utility of the ZI-OU model, we fitted a traditional OU model without regime switching to the simulated data with regime switching. The posterior predictive trajectories, as shown in the lower left panel of Fig. 5, indicated that the OU model cannot capture the disparate trajectories implicated, respectively, in the active and inactive regimes, and resulted in wide predictive bands. Moreover, to examine whether the ZI-OU would over-fit the data if no regime switching existed, we simulated data from a traditional OU model using the setting of the active regime of the ZI-OU model in the simulation, and then we fit a ZI-OU model to the data. The posterior predictive trajectories are shown in the lower right panel of Fig. 5. The result demonstrated that the ZI-OU model correctly classified the trajectories of the OU model as being in the active regime. In sum, the ZI-OU model does not exhibit over-fitting problems or indicate the existence of a spurious regime when the true model is a single-regime model.

6. Empirical Illustration

Our motivating example was built on a longitudinal study of young children's self-regulation during the frustration-inducing transparent lock box task at 36 and 48 months of age. The lock box task lasted approximately 150 s (2.5 min) at 36 months, following the standard protocol from Lab-TAB (Goldsmith & Reilly, 1993). For the same children at 48 months, the task lasted 300 s: It consisted of two parts to increase the probability of witnessing PR-related behaviors in the older children. Specifically, the first 180 s (3 min) of the task was administered in identical ways to the procedures for children at 36 months, except that the child was left with the lock box for 30 s longer. After 180 s, a research assistant entered the room and told the child to keep opening the lock box, after which the research assistant left and returned after 120 s (making the task last a total of 300 s or 5 min).

We used composite scores of PR and EP obtained from aggregating multiple binary markers in the video-coded data at each observed time point over the durations of task, yielding approximately $T = 150$ and 300 measurement occasions for each of the $n = 128$ and 119 children at 36 and 48 months, respectively. Behavioral markers of PR were operationally defined by children's anger expression in face or voice, angry verbalization about the problem, sadness in face or voice, and other signs of tension in face or voice. Behavioral markers of EP were operationally defined by self-soothing, attempt to open the box with the key but not engaging in any disruptive act, and attempt to open the box a in different way appropriately (e.g., see if it can open at the hinge).

6.1. Data Processing Details

Several preliminary data screening procedures were performed prior to model fitting. First, children with excessively high amounts of missing data (with $T_i < 20$) due to technical difficulties (e.g., they were out of the camera) were excluded from model fitting. Second, we also excluded children with excessively large chunks of consecutively missing (> 20) time points to avoid biasing the parameter estimates in the "active" regime of the OU process. If the number of consecutively missing time points was less than 20, the participants were retained and the missing time points were estimated as latent variables in the MCMC algorithm. Finally, children with less than 100

time points due to early termination of the experimental procedures were excluded from the analysis. The data screening procedures yielded a final sample size of $n = 213$ (111 children at 36 months) for final model fitting, with total number of time points ranging from 103 to 150 for children at 36 months, and 117–300 for children at 48 months. We smoothed the trajectories by calculating the moving averages with a window of ± 3 s.

6.2. Model Reduction

We fit the proposed model summarized in Eqs. (2)–(4) to the empirical data. The prior distributions in (7) and (8) are used, and further details are given later after the logistic model (4) is specified. Preliminary inspection of the trace plots indicates that the OU-related parameters, $\beta_{j,1}$, for the EP and PR processes were not empirically identifiable. These specific parameters affect how rapidly the EP and PR processes approach zero once they are in the inactive regime. Because the reduction to zero activity occurred almost instantaneously in the empirical data, there was insufficient information from the data to allow these parameters to be estimated uniquely. This was in contrast to results from the above simulation study in which the transition unfolded over a longer duration, and $\beta_{j,1}$ could indeed be uniquely estimated. Thus, for the empirical data we set these parameters to $\beta_{j,1} = 60$, $\mu_{j,1} = 0$, and $\sigma_{j,1} = 0.01$ for both EP and PR to yield relatively rapid rates of approaching zero once these processes transitioned into the inactive regime and relatively little fluctuations around zero once they settled into the zero activity period.

As noted earlier, for identification purposes, we set the log-odds terms dictating a change of regime for any of the two processes (EP and PR) to zero regardless of the previous regime of the opposing process. Thus, $\alpha_{121}^{\text{EP}}, \alpha_{211}^{\text{EP}}, \alpha_{122}^{\text{EP}}, \alpha_{212}^{\text{EP}}, \alpha_{112}^{\text{PR}}, \alpha_{211}^{\text{PR}}, \alpha_{122}^{\text{PR}}$, and α_{221}^{PR} were all set to $\mathbf{0}$, a vector of zeros of appropriate dimension. In addition to the minimum identification constraints, we used a forward–backward approach to simplify the structure of the age-specific regime-switching (i.e., Markov transition) functions. To do so, we first fit the proposed model without any age effects to the data from both age-groups. Thus, \mathbf{u}_i consisted only of a unit constant to define the intercept terms. The coefficients of the LOs are listed in Table 1 under “full main effects.” We checked if the coefficients for the transitional probabilities of one process given the two regimes of the opposing process at the current time point are substantially different. If the 90% credible interval of the difference of the coefficients covered zero, we constrained the pair of coefficients to be identical. For example, $\alpha_{212,0}^{\text{PR}}$ and $\alpha_{222,0}^{\text{PR}}$ are not substantially different and are replaced by a common parameter denoted by $\alpha_{2,2,0}^{\text{PR}}$. Second, we added age effects to the model in addition to the constrained main effects, which is listed in the upper part of Table 2 labeled “full age effects.” The age effects with 90% credible intervals covering zero are removed. Moreover, if the difference of the age coefficients for one process given the two state of the opposing process at the current time point is not substantial different (its 90% credible interval included zero), the pair of coefficients are constrained to be identical. Finally, the main effects are checked with the procedure in the first step in the presence of the selected age effects. The resulting parameters of LO are shown in the lower part of Table 2 termed “reduced main and age effects.” The constrained parameters are those with “.” in the subscript.

Some pairs of constrained coefficients set to be identical warrant some clarifications here. Specifically, when the model with all age-related effects incorporated was fitted to data from both age-groups, we found no substantial difference in the LO for the 36-month children’s PR (i.e., the age-group coded as 0 on the covariate, age_i) to stay inactive regardless of whether their EP was previously active or inactive. Thus, we set $\alpha_{111,0}^{\text{PR}}$ to be equal to $\alpha_{121,0}^{\text{PR}}$ and simplified the corresponding notation to $\alpha_{1,1,0}^{\text{PR}}$ in subsequent models. In doing so, the intercepts associated with the transition probability for PR to stay within the inactive regime from the current to the next time were constrained to be the same regardless of whether the previous regime for EP was active or inactive. In a similar vein, the probability for PR to stay active also did not depend reliably on

the previous regime of EP, thus motivating us to set $\alpha_{212,0}^{PR}$ to be equal to $\alpha_{222,0}^{PR}$ and simplify the corresponding notation to $\alpha_{2,2,0}^{PR}$. Other age-related LO parameters whose 90% credible intervals included zero were also removed, and the pairs of age-related coefficients were set to be identical following the rule for the intercepts. After pruning the age-related coefficients, the main effects are checked for redundancy again. The $\alpha_{111,0}^{PR}$ and $\alpha_{121,0}^{PR}$ were not substantially different and constrained to $\alpha_{1,1,0}^{PR}$ in the lower part of Table 2.

6.3. *Reduced Regime-Switching Functions*

After removing all unimportant coefficients and all coefficients that were set to zero for identification purposes and constraining similar parameters to be identical, we obtained via a simplified model:

Conditional probabilities of EP to be inactive or “off”

$$p\left(l_{i,k}^{EP} = 1 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 1\right) = \frac{\exp\left\{\alpha_{111,0}^{EP}\right\}}{1 + \exp\left\{\alpha_{111,0}^{EP}\right\}}, \tag{9}$$

$$p\left(l_{i,k+1}^{EP} = 1 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 2\right) = \frac{\exp\left\{\alpha_{112,0}^{EP}\right\}}{1 + \exp\left\{\alpha_{112,0}^{EP}\right\}}, \tag{10}$$

where, as alluded earlier, $p\left(l_{i,k}^{EP} = 2 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 1\right) = 1 - p\left(l_{i,k}^{EP} = 1 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 1\right)$, and $p\left(l_{i,k+1}^{EP} = 2 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 2\right) = 1 - p\left(l_{i,k+1}^{EP} = 1 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 2\right)$ by nature of the identification constraints. Similarly, we obtained:

Conditional probabilities of EP to be active or “on”

$$p\left(l_{i,k+1}^{EP} = 2 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 1\right) = \frac{\exp\left\{\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP} \text{age}_i\right\}}{1 + \exp\left\{\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP} \text{age}_i\right\}}, \tag{11}$$

$$p\left(l_{i,k+1}^{EP} = 2 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 2\right) = \frac{\exp\left\{\alpha_{222,0}^{EP}\right\}}{1 + \exp\left\{\alpha_{222,0}^{EP}\right\}}, \tag{12}$$

whereas $p\left(l_{i,k+1}^{EP} = 1 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 1\right) = 1 - p\left(l_{i,k+1}^{EP} = 2 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 1\right)$, and $p\left(l_{i,k+1}^{EP} = 1 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 2\right) = 1 - p\left(l_{i,k+1}^{EP} = 2 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 2\right)$ due, again, to the identification constraints. Similar conditional probabilities were estimated for PR, though with slightly different “important” age-related effects as:

Conditional probabilities of PR to be inactive or “off”

$$p\left(l_{i,k+1}^{PR} = 1 | l_{i,k}^{EP} = 1, l_{i,k}^{PR} = 1\right) = \frac{\exp\left\{\alpha_{1,1,0}^{PR} + \alpha_{111,1}^{PR} \text{age}_i\right\}}{1 + \exp\left\{\alpha_{1,1,0}^{PR} + \alpha_{111,1}^{PR} \text{age}_i\right\}}, \tag{13}$$

$$p\left(l_{i,k+1}^{PR} = 1 | l_{i,k}^{EP} = 2, l_{i,k}^{PR} = 1\right) = \frac{\exp\left\{\alpha_{1,1,0}^{PR} + \alpha_{121,1}^{PR} \text{age}_i\right\}}{1 + \exp\left\{\alpha_{1,1,0}^{PR} + \alpha_{121,1}^{PR} \text{age}_i\right\}}, \tag{14}$$

Conditional probabilities of PR to be active or “on”

$$p\left(l_{t_i,k+1}^{\text{PR}} = 2 | l_{t_i,k}^{\text{EP}} = 1, l_{t_i,k}^{\text{PR}} = 2\right) = \frac{\exp\left\{\alpha_{2,2,0}^{\text{PR}}\right\}}{1 + \exp\left\{\alpha_{2,2,0}^{\text{PR}}\right\}}, \quad (15)$$

$$p\left(l_{t_i,k+1}^{\text{PR}} = 2 | l_{t_i,k}^{\text{EP}} = 2, l_{t_i,k}^{\text{PR}} = 2\right) = \frac{\exp\left\{\alpha_{2,2,0}^{\text{PR}}\right\}}{1 + \exp\left\{\alpha_{2,2,0}^{\text{PR}}\right\}}, \quad (16)$$

where our choice of identification constraints, again, entailed the following conditional probabilities that could just be computed from other modeling parameters, as opposed to being estimated, including: $p(l_{t_i,k+1}^{\text{PR}} = 2 | l_{t_i,k}^{\text{EP}} = 1, l_{t_i,k}^{\text{PR}} = 1)$, $p(l_{t_i,k+1}^{\text{PR}} = 2 | l_{t_i,k}^{\text{EP}} = 2, l_{t_i,k}^{\text{PR}} = 1)$, $p(l_{t_i,k+1}^{\text{PR}} = 1 | l_{t_i,k}^{\text{EP}} = 1, l_{t_i,k}^{\text{PR}} = 2)$, and $p(l_{t_i,k+1}^{\text{PR}} = 1 | l_{t_i,k}^{\text{EP}} = 2, l_{t_i,k}^{\text{PR}} = 2)$.

All intercept terms—namely parameters whose fourth digit in the subscript takes on a 0 (e.g., $\alpha_{111,0}^{\text{EP}}$)—capture the LOs for children at 36 months to show the same EP and PR statuses at the next as well as the current time point, relative to showing a switch in EP and PR statuses. All age-related parameters—namely, parameters whose fourth digit in the subscript takes on a 1 (e.g., $\alpha_{221,1}^{\text{EP}}$)—capture the deviations in LOs for children at 48 months, compared to those at 36 months, to show the same EP and PR statuses at the next as well as the current time point, relative to showing a switch in EP and PR statuses. As shown in (9)–(16), only three of such age-related effects were determined to be reliably different from 0, including: $\alpha_{221,1}^{\text{EP}}$, $\alpha_{111,1}^{\text{PR}}$, and $\alpha_{121,1}^{\text{PR}}$. These parameters are referred to herein as “*direct parameters*” because they were present directly in the reduced model. This was in contrast to other “*derived parameters*” that we derived from the direct parameters to answer targeted theoretical questions of interest, as elaborated in detail in the context of the modeling results.

Diffuse prior distributions were used for the parameters in (7) and (16). Specifically, the means of the prior distributions of $\beta_{\text{EP},2}$ and $\beta_{\text{PR},2}$ were set equal to 0.2, and those of $\mu_{\text{EP},2}$ and $\mu_{\text{PR},2}$ were set to 2. Their prior variances were set to 10^2 . The hyperparameters of the diffusion parameters were set as $a_{\text{EP},12} = a_{\text{PR},12} = 2$, $a_{\text{EP},22} = a_{\text{PR},22} = 0.01$. A diffuse univariate normal distribution with mean of 0 and variance 10^4 was assigned to the coefficients in (9)–(16).

6.4. Results from the Final Model

We first describe the SDE-related estimates within the active regime, followed by elaborations on results pertaining to age differences in regime-switching properties. The point and standard error estimates of the parameters in the SDE functions while individuals were in the active regime are shown in Table 4. These parameters include the equilibria ($\mu_{\text{EP},2}$ and $\mu_{\text{PR},2}$), the approach rates ($\beta_{\text{EP},2}$ and $\beta_{\text{PR},2}$), and the diffusion parameters ($\sigma_{\text{EP},2}$ and $\sigma_{\text{PR},2}$) associated with EP and PR. The estimated equilibria of EP and PR were similar in value compared to each other (around 1.00) and were sufficiently distinct from zero, thus providing some face validity to the definition of the second regime as the active regime. We found that EP was characterized by a slightly larger approach rate than the PR, indicating that when EP switched from the inactive regime to the active regime, it approached its equilibrium more quickly compared to PR. In comparison, the diffusion parameter of EP was smaller in magnitude than that of PR, suggesting that EP was more stable than PR—or showed less variability around its equilibrium—while in the active regime. Collectively, these parameters help shed light on the nature and intrinsic dynamics of EP and PR in young children when these processes are active.

We organized our regime-switching results in groups, elaborating first on results pertaining to (1) the direct parameters in our reduced model, followed by results based on: (2) derived

TABLE 4.
The point (Est) and standard errors (SE) estimates of the SDE parameters while in the active regime.

Parameters	Est	SE	Parameters	Est	SE
$\beta_{EP,2}$	2.792	0.089	$\beta_{PR,2}$	2.075	0.099
$\mu_{EP,2}$	1.038	0.014	$\mu_{PR,2}$	0.962	0.033
$\sigma_{EP,2}^2$	0.746	0.006	$\sigma_{PR,2}^2$	0.860	0.010

Est, estimates with the means of MCMC samples from the posterior distribution; SE, standard error estimated by the standard deviation of MCMC samples from the posterior distribution.

parameters for 36-month children; (3) derived parameters for 48-month children; and (4) age differences in the derived parameters. We summarize in Tables 5 and 6 these parameters' estimated posterior means, standard deviations (labeled as SD), credible interval, and the pseudo p values (labeled as “ p value”) calculated with the posterior distribution analogue to the frequentist p value based on sampling distributions. The posterior distributions of the parameters are close to normal distribution, and we used normal distribution to compute these pseudo p values to quantify the distance of 0 from the posterior means.

6.4.1. (1) Results Pertaining to the Direct Parameters (Set 1) The estimates for the direct parameters provided some initial glimpses into the dependencies between EP and PR and age differences therein. Specifically, the direct parameters, $\alpha_{221,1}^{EP}$, $\alpha_{111,1}^{PR}$, and $\alpha_{121,1}^{PR}$, revealed whether children at 48 months, when compared to themselves at 36 months, were more adept at keeping EP “on” when PR was previously “off” (see effect 1a in Table 5, linked to parameter $\alpha_{221,1}^{EP}$) and at keeping PR “off” when EP was previously off and on (effects 1b and 1c). Based on the direct age-related effects, the following developmental shifts in transition probabilities between the 36-month and 48-month children can be noted. First, the probability for EP to remain active given that PR was inactive at the previous time was lower (i.e., the log-odd parameter, $\alpha_{221,1}^{EP} = -0.229$, was negative) for the older than the younger age-group. In other words, EP was less likely to stay active among the older children when PR was previously inactive. This may suggest that the 48-month children were quicker or more effective at deploying EP-related strategies only as needed, namely only turning on EP when it was most needed. Second, the probabilities for PR to remain inactive when EP was previously active ($\alpha_{121,1}^{PR} = -0.878$) or inactive ($\alpha_{111,1}^{PR} = -0.358$) were both lower for the 48-month children compared to the 36-month children. This suggested that contrary to the theoretical postulate that PR may be lower in older than younger children due to increased self-regulation skills, older children were actually more likely to show activation in PR than younger children regardless of the status of EP.

6.4.2. Results Based on the Derived Parameters (Sets 2–4) Using estimates for the direct parameters, we then computed the derived parameters in sets (2) and (3) to indicate differences in the lagged (previous) effects of the opposing process on the transition LOs of a process when the opposing process was previously “on” versus “off” for each one of the two age-groups. Note that the effects summarized under sets 2 and 3 were simply proxies to help us calculate the age difference effects in set 4, which capture the key essence of self-regulation as we originally conceptualized. In particular, we were interested in testing whether the older age-group showed more substantial differences in the LOs of sustaining EP between PR was previously “on” and “off” (effect 4a); and whether having EP as previously “on” as opposed on “off” facilitated efforts to keep PR “off” (effect 4b). The former (effect 4a) may be regarded as a kind of age difference in the ability to *counter the regulatory interference of PR on EP* when PR was “on”

TABLE 5.
Theoretical concepts of interest in our motivating example, parameters capturing these concepts, and the related equations.

(1) Direct parameters	Par	Equations	Mean	SD	CI	<i>p</i> value
(1a) Age difference of LO of EP staying active given PR _{<i>t_{i,k}</i>} inactive	$\alpha_{221,1}^{EP}$	(11)	-0.229	0.116	[-0.421, -0.038]	0.049
(1b) Age difference of LO of PR staying inactive given EP _{<i>t_{i,k}</i>} inactive	$\alpha_{111,1}^{PR}$	(13)	-0.358	0.095	[-0.512, -0.201]	< 0.001
(1c) Age difference of LO of PR staying inactive given EP _{<i>t_{i,k}</i>} active	$\alpha_{121,1}^{PR}$	(14)	-0.878	0.08	[-1.003, -0.742]	< 10 ⁻²⁷
(1d) LO of EP staying inactive given PR _{<i>t_{i,k}</i>} inactive in 36 months	$\alpha_{111,0}^{EP}$	(9)	3.49	0.057	[3.40, 3.59]	
(1e) LO of EP staying inactive given PR _{<i>t_{i,k}</i>} active in 36 months	$\alpha_{112,0}^{EP}$	(10)	3.23	0.078	[3.10, 3.36]	
(1f) LO of EP staying active given PR _{<i>t_{i,k}</i>} inactive in 36 months	$\alpha_{221,0}^{EP}$	(11)	4.17	0.092	[4.02, 4.32]	
(1g) LO of EP staying active given PR _{<i>t_{i,k}</i>} active in 36 months	$\alpha_{222,0}^{EP}$	(12)	3.83	0.066	[3.72, 3.95]	
(1h) LO of PR staying inactive given EP _{<i>t_{i,k}</i>} inactive or active in 36 months	$\alpha_{1,1,0}^{PR}$	(13), (14)	3.91	0.067	[3.80, 4.02]	
(1i) LO of PR staying active given EP _{<i>t_{i,k}</i>} inactive or active in 36 months	$\alpha_{2,2,0}^{PR}$	(15), (16)	2.70	0.034	[2.65, 2.76]	

LO, log-odds; CI, credible interval.

TABLE 6.
Theoretical concepts of interest in our motivating example, parameters capturing these concepts, and the related equations.

<i>(2) Derived parameters for the 36 months group</i>						
(2a) Difference in the LO of EP switching from inactive to active when $PR_{t_i,k}$ was active versus inactive	$\alpha_{111,0}^{EP} - \alpha_{112,0}^{EP}$	(9), (10)	0.257	0.098	[0.093, 0.417]	0.009
(2b) Difference in the LO of EP staying active when $PR_{t_i,k}$ was active versus inactive	$\alpha_{222,0}^{EP} - \alpha_{221,0}^{EP}$	(11), (12)	-0.344	0.113	[-0.528, -0.160]	0.002
<i>(3) Derived parameters for the 48 months group</i>						
(3a) Difference in the LO of EP switching from inactive to active when $PR_{t_i,k}$ was active versus inactive	$\alpha_{111,0}^{EP} - \alpha_{112,0}^{EP}$	(9), (10)	0.257	0.098	[0.093, 0.417]	0.009
(3b) Difference in the LO of EP staying active when $PR_{t_i,k}$ was active versus inactive	$\alpha_{222,0}^{EP} - (\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP})$	(11), (12)	-0.115	0.095	[-0.274, 0.042]	0.229
(3c) Difference in the LO of PR staying inactive when $EP_{t_i,k}$ was active versus inactive	$\alpha_{121,1}^{PR} - \alpha_{111,1}^{PR}$	(13), (14)	-0.52	0.082	[-0.657, -0.386]	$< 10^{-9}$
<i>(4) Age differences in the derived parameters</i>						
(4a) Age effect on the difference in the LO of EP staying active when $PR_{t_i,k}$ was active versus inactive	$-\alpha_{221,1}^{EP}$	(11)	0.229	0.116	[0.038, 0.421]	0.049
(4b) Age effect on the difference in LO of PR staying inactive when $EP_{t_i,k}$ was active versus inactive	$\alpha_{121,1}^{PR} - \alpha_{111,1}^{PR}$	(13), (14)	-0.52	0.082	[-0.657, -0.386]	$< 10^{-9}$

LO, log-odds, CI, credible interval. The effects at 36 months similar to (3c) are also of interest. However, the corresponding function of parameters, $\alpha_{121,0}^{PR} - \alpha_{111,0}^{PR}$ was not reliably different from zero based on the model with “full age effects” model in Table 2 (Est = 0.075, SD = 0.149, 90% CI = [-0.178, 0.316]), and thus the two parameters are constrained to be identical as $\alpha_{1,1,0}^{PR}$ and the difference is fixed at zero.

versus “off,” whereas the latter (effect 4b) may be understood as a way to test the hypothesis of *increased regulatory efficacy of EP* with age—namely older children’s ability to show increased likelihood of keeping PR “off” when EP was previously “on” as opposed to “off.” These effects were not directly observed effects attributable to a single parameter, but rather involved functions of several modeling parameters. As such, it is possible, but somewhat cumbersome, to compute the standard errors and uncertainties associated with these derived parameters (via delta method in the frequentist framework). This process can be accomplished with minimal added computational costs using MCMC samples from the Bayesian estimation procedures.

Of the derived parameters in sets (2) and (3)—at 36 months and 48 months, respectively—we did not find any age difference in the effect of previous PR in triggering EP (from “off” to “on”) when PR was previously active versus inactive (see effects 2a and 3a in Table 6). For both age-groups, the derived parameter, $\alpha_{111,0}^{EP} - \alpha_{112,0}^{EP} = 0.257$, suggested that active PR at the previous time point was more likely to invoke or trigger EP, and there was no statistically notable age difference in the ability to invoke/deploy EP when needed.

Comparisons of the estimates for the derived parameter in (2b) and (3b) for the two age-groups did suggest some age difference in the predicted LO for EP to remain active when PR was previously active versus inactive. In particular, EP was substantially less likely to stay active when PR was previously active at 36 months (see effect 2b; as given by $\alpha_{222,0}^{EP} - \alpha_{221,0}^{EP} = -0.344$). A similar tendency was observed for children at 48 months, even though the difference was associated with a credible interval that included 0 (effect 3b, $\alpha_{222,0}^{EP} - (\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP}) = -0.115$). That is, both age-groups showed some evidence of *regulatory interference* of PR on EP.

Consistent with our expectation, the older age-group, when compared to the 36-month children, actually showed a *smaller* difference between LOs of maintaining active EP when PR was previously active versus inactive. This age difference, as given by $[\alpha_{222,0}^{EP} - (\alpha_{221,0}^{EP} + \alpha_{221,1}^{EP})] - [\alpha_{222,0}^{EP} - \alpha_{221,0}^{EP}] = -\alpha_{221,1}^{EP} = 0.229$, was reliably different from zero, as tested explicitly under (4a). In other words, the EP of the 48-month group did not show as much of a change in LO of staying active regardless of whether PR was previously active or inactive and was less susceptible to the regulatory interference of PR than the 36-month children. The tendency to sustain EP (which included markers such as engaging in “on-task” behaviors) regardless of the status of PR is a direct reflection of the older children’s more mature self-regulation skills compared to the younger children.

Finally, examination of age differences in *regulatory efficacy* as conveyed by the lagged effects from EP on PR revealed the followings. First, the probability of PR staying inactive when EP was previously active versus inactive did not differ reliably for the 36-month children. In other words, the corresponding derived parameter for the younger age-group, given by $\alpha_{121,0}^{PR} - \alpha_{111,0}^{PR}$ (Est = 0.075, SD = 0.149, 90% CI = [-0.178, 0.316]), was not reliably different from zero based on the model with “full age effects” in Table 2. Consequently, $\alpha_{121,0}^{PR}$ and $\alpha_{111,0}^{PR}$ are set to be identical as $\alpha_{1.1,0}^{PR}$ in the final model. This difference in log-odds was substantially lower than zero at 48 months (see effect 3c; $\alpha_{1.1,0}^{PR} + \alpha_{121,1}^{PR} - (\alpha_{1.1,0}^{PR} + \alpha_{111,1}^{PR}) = (\alpha_{121,1}^{PR} - \alpha_{111,1}^{PR}) = -0.52$). Computing the age difference in this effect under (4b) confirmed that this age difference was indeed reliably different from zero. That is, contrary to our expectation, we found that older children showed reduced *regulatory efficacy of EP*—namely lower likelihood of keeping PR inactive when EP was previously active as opposed to inactive. Thus, the older age-group was actually less able to desist PR given previous activation of EP, thereby showing more instances where PR and EP were both active.

Model-implied trajectories generated using the final model are plotted in Fig. 3 for one hypothetical child at 36 (see the upper panel) and 48 (see the lower panel) months, respectively, generated using parameters obtained from model fitting. In concordance with the findings described earlier, the hypothetical subject at 36 months is observed to show more EP deactiva-

tion when PR is active versus inactive, consistent with the finding that at 36 months, the children in our sample were more susceptible to regulatory interference of PR on EP than they were at 48 months. However, there is an absence of previous studies evaluating the effectiveness of such EP-related strategy use (for an exception see Buss & Goldsmith, 1998). We found—contrary to the theoretical postulate that increased use of EP-related strategy would be associated with decreased PR—decreased regulatory efficacy of EP, or in other words, increases in PR at 48 than at 36 months. Thus, the hypothetical subject in Fig. 3 is also observed to show more instances of simultaneous co-activation of PR and EP (i.e., shown as overlap between the stacked upper and lower shaded regions). The increased tendency for such co-activation of PR and EP at the age of 48 months than at 36 months may mean that despite increased use of regulatory strategies, the 48-month children were more frustrated during the task than the 36-month children.

7. Discussion

In this paper, we proposed a ZI-OU model motivated by a laboratory study of children's self-regulatory behaviors, in which the occurrence of the behaviors of interest was observed to be interspersed with extended, consecutive periods of non-occurrence (zero responses). Building on our motivating example, we presented a broader regime-switching SDE modeling framework within which the ZI-OU can be viewed as a special case. Bayesian estimation and inference were developed for the general regime-switching SDE framework and applied to the ZI-OU model as a special case. The performance of these estimation and inferential procedures was evaluated with a simulation study, and their practical utility was demonstrated with the motivating empirical data.

7.1. Summary and Implications of the Empirical Results

7.1.1. Summary In the present study, we used a Markov-based regime-switching SDE model to represent the dynamic interdependencies between two processes—EP and PR—in the context of children's self-regulation processes. Using the proposed model, we found that for both age-groups, EP was less likely to stay “on” when PR was previously active, but the 48-month children, when compared to those at 36 months, showed a significantly smaller difference in log-odds of maintaining active EP between instances with active and inactive previous PR. In addition, the older age-group also showed statistically significant increase in the *regulatory efficacy of EP*—namely increased likelihood of EP keeping PR in the inactive regime when EP was “on” as opposed to “off.” The OU model has been used to represent individuals' emotion regulatory dynamics due to its ability to capture homeostatic dynamics as exponential return to a baseline (e.g., Oravec et al., 2011). The results illustrate the value of constructing regime-switching extensions to the traditional OU model to enable simultaneous accommodation of postulates of affective dynamics and identification of measurement regimes (active and inactive) that are vital for understanding children's development of the ability to manage negative emotions. Generally, results align with theoretical propositions about age-related increases in self-regulation ability, but also suggest that nuances in the dynamic relation between PR and EP are more complicated than hypothesized. In this sense, the proposed model pushes for more detailed theory about exactly what aspects of the interplay between EP and PR are marking children's developing self-regulation ability.

7.1.2. Linkages to Previous Self-Regulation Studies It may be worth noting that we previously developed and utilized a multilevel, nonlinear ordinary differential equation model (Chow, Bendezú, Cole, & Ram, 2016; Cole, Bendezú, Ram, & Chow, 2017) to capture regulatory interference of PR on EP among 36-month children—operationalized as the effect of PR in damping,

or reducing the over-time amplitude of EP's oscillatory dynamics, and regulatory efficacy—operationalized as the effect of EP in damping the over-time amplitude of PR' oscillatory dynamics. In the ordinary differential equation (ODE) context, these phenomena were operationalized as parameters in the nonlinear ODE model. In the present context, they were represented within a regime-switching SDE model by means of transition log-odds parameters. In addition, the previous ODE model uses second derivatives as dependent variables in the ODE function and, as such, it was assumed that EP and PR would show oscillatory trajectories even in the absence of interactive effects with the other process, and that the counteracting effects of EP and PR on each other would be manifested as over-time reductions in these processes' oscillatory dynamics. Such oscillatory nature of the PR and EP processes was more evident in the data used in our previous ODE modeling, which involved the use of a different task for eliciting PR-related responses than the lock box task used in our current sample.

Thus, in the current regime-switching SDE model, we instead used the OU model—a model with first derivatives as the dependent variables—in the non-ZI portion of the model. A direct consequence of doing so is that EP and PR were no longer postulated to show self-sustaining oscillatory dynamics, and the counteracting effects between EP and PR were relegated, instead, to the regime switching functions, as opposed to being incorporated into the SDE functions. Compared to the previously proposed ODE model, the new proposed model can better account for the ZI characteristic of the data—a feature left unmodeled in the ODE model used in our previous work. Furthermore, the proposed regime-switching SDE model also extended our previous ODE model by using SDEs to capture the stochasticity (process noises) in the dynamic functions. Nevertheless, the shift from showing oscillatory dynamics in PR and EP to more structured fluctuations around one's EP and PR baselines may represent a developmental shift that is worth further investigation in future studies.

One related variation to the empirical model proposed herein is a ZI Poisson model as the measurement model (Roeder, Lynch, & Nagin, 1999) and an SDE to represent over-time changes in the intensity rate of the Poisson process in the non-ZI portion of the data. We did not pursue this option in the present context due to our interest in explicitly modeling the transition between the inactive and active regimes of EP and PR as dependent on each other, and the lack of sufficient “spread” in the nonzero counts observed in the current data set (i.e., with most nonzero counts clustering around 1 and 2) to reliably distinguish the intensity rate of the Poisson process from the zero counts serving to identify the inactive regime. It would be interesting to do a more targeted comparison of these two related modeling variations using data sets that do show more variability in the nonzero counts.

7.2. Summary and Implications of the Simulation Study

The simulation study was designed to test the performance of the estimation procedures under conditions that mirrored the settings and characteristics of the observed data. The estimates of the model parameters and the latent regime indicators are accurate. The Bayesian results are not sensitive to the two alternative informative prior settings. The convergence of the MCMC algorithm is fast given reasonable starting values. The posterior predictive trajectories demonstrate the excellent model fit of the ZI-OU model applied to data with or without the presence of zero-inflated intervals. In comparison, the classic OU model leads to poor model fit for the data with considerable zero-inflated intervals.

7.3. Limitations, Unresolved Challenges, and Future Extensions

7.3.1. Model Setup In this article, we considered a specialized model designed to capture selected characteristics of young children's self-regulation processes. The simulation study was designed to test the performance of the estimation procedures under conditions that mirrored

the settings and characteristics of the observed data. As such, the simulation study is, understandably, limited in generalizability to other settings and data conditions. Beyond broadening the conditions considered to include a wider range of parameter values, effect sizes, and sample size configurations, simulations involving other variations of the proposed model are also worth pursuing in future studies. For instance, possible extensions may include incorporating nonlinear ODE or SDE functions (Chow, Bendezú, et al., 2016 ; Lu et al., 2015) into the active regime of the proposed model, as well as adding other observed predictors and interactions terms in the RS model to represent context-specific transition probabilities.

A clear empirical limitation in the present example is that the lock box task used to generate EP- and PR-related behaviors varied slightly in design at the two ages (i.e., longer at 48 months and with additional “intervention” from the experimenter to remind the 48-month children of the potential reward from the task). Even though the coding system used to code the behaviors associated with the two instances of lock box task is identical at the two ages and thus helped enforce some measurement invariance constraints, any conclusions concerning age differences reported in this present article could, in principle, arise from differences in the nature of the administered task and thus have to be interpreted with caution. Moreover, we treated the data at 36 and 48 months as two independent age-groups even though these were within-subjects data that might show other added sources of serial dependencies over time. The limited number of repeated assessments (over two ages) has limited us from considering other more complex age-based structures and inter-individual differences over ages. With more within-participant assessments over multiple ages, we may be able to add random effects to the proposed regime-switching SDE model.

The empirical application in the present article was motivated by ZI that arises from tallying empirically coded instances of specific behaviors into composite scores representing EP and PR processes. As such, we imposed, on the basis of theoretical grounds, a “measurement model” that derives two set of scores (corresponding to EP and PR) for each individual and time point through predetermined weights. Future studies should formally evaluate the tenability of this heuristic measurement approach, and whether measurement invariance indeed holds across ages.

7.3.2. Estimation One novel feature of the proposed model was the dependency of the regime-switching process not only on one but two regime indicators (i.e., the regime indicators for PR and EP). This modeling extension was motivated by our interest in representing the dynamic interdependencies between EP and PR in the context of self-regulation. However, incorporating two regime indicators into the model was not without a cost: Doing so expanded the possible cases (or cells) in the transition probability matrix from $2^2 = 4$ to $2^4 = 16$. The number of parameters in the logistic model characterizing the transitional probabilities also increased accordingly. Consequently, more computational time was required, and we encountered difficulties in estimating the parameters in some of these cells due to the lack of sufficient observations to empirically identify these parameters. We resorted to simplifying our model for transition log-odds in a piecemeal way given the moderate sample size we had. Larger sample sizes would be needed to provide enough observations and instances of transition through the various cells to fit the full model.

Another estimation difficulty that arose in the empirical study but not the simulation study was that particular parameters, such as the approach rates for EP and PR during the inactive regime, were not empirically identifiable but could be estimated in the simulation study. The reason was that the rates at which EP and PR approached zero during the inactive regime were too fast (basically instantaneously) in the empirical study due to the excessive preponderance of “zero counts” in the data. Our coding of the raw empirical data was already performed at the second-by-second level so further increase in the density of measurements between the second-by-second intervals is not viable from a practical standpoint. Moreover, to keep the dynamic model parsimonious, we assumed the equilibrium of each process within a regime is the same

across children. However, a viable alternative is to increase the length of the tasks that generate the EP- and PR-related behaviors to yield longer sequences of data for modeling purposes, and to increase the number of markers for measuring the EP and PR processes. Moreover, informative prior distribution can be adopted to regularize the estimation of these parameters. However, elicitation of prior information may be challenging.

From an estimation standpoint, we utilized the default MCMC algorithms in JAGS for modeling purposes. These algorithms were not designed to handle the specific model proposed here. Some methodological adaptations of these algorithms are thus possible to improve the estimation properties of these algorithms, particular with regard to computational efficiency. For instance, when the number of time points is large, updating estimates for the latent regime indicator for each person and time point entailed prohibitively long computational time—an unfortunate cost given that the children in this study showed high stability of staying within the inactive regime once they transition into it and, consequently, it may not be necessary to update the estimates of the latent regime indicator at every single time point. Considerable improvements in computational time may be achieved by developing a modified MCMC algorithm that can adaptively determine the time points at which updates of the latent regime indicator are required. In addition, in other applications involving large intervals between consecutive observations, imputation of additional time points between adjacent observed time points as required by the Euler approximation may lead to substantial reduction in computational efficiency. To this end, advanced sampling methods (Durham & Gallant, 2002; Lindström, 2012; Lu et al., 2015) may be used instead to improve the efficiency of the default MCMC algorithms implemented in standard Bayesian software packages. Alternatively, the model could also be handled in a frequentist framework through stochastic expectation maximization algorithms, where the regime indicators are latent variables and the maximum of parameters are calculated iteratively. (Chow, Lu, Sherwood, & Zhu, 2016; Solo, 2002)

7.3.3. Inference In the ZI-OU, the meanings of the regimes (i.e., active and inactive) are predefined based on theoretical expectations. Hence, the number of regimes is also predetermined. In this study, we compared the fit of the ZI-OU model relative to that of the traditional one-regime OU model by investigating the predictive trajectories. Other model comparison criteria, such as the deviance information criteria (DIC, Spiegelhalter, Best, Carlin, & Van Der Linde, 2002) and Bayesian leave-one-out approach with Pareto smoothed importance sampling (LOO-PSIS Vehtari, Gelman, & Gabry, 2016) may also be used in other situations involving general regime-switching SDEs. However, in the presence of latent variables and SDEs without analytical solutions, it is challenging to apply these existing Bayesian model comparison criteria. Their performance in determining the correct number of components warrants more extensive future studies.

We used credible intervals and pseudo p values to “test” if the parameters or their functions are substantially different from 0. Although they look similar to the confidence intervals and p values in frequentist approaches, the interpretations are different. Instead of the sampling distribution under the null hypothesis, the Bayesian analogues are based on the posterior distribution of the parameters of interest and measure the distance of 0 from the center of the posterior distribution. Another approach that is similar to frequentist approaches in the sense of comparing a null hypothesis with an alternative hypothesis is using Bayes factor (Kass & Raftery, 1995). However, its computational time is considerable, especially for complex models. Our existing approach using credible intervals and pseudo p values provides a viable way to address the importance of the parameters in the model.

7.4. Closing Remarks

In this paper, we modeled age differences in the second-to-second progression of children’s self-regulation using an SDE model—specifically a bivariate OU model. Data constraints inherent

in observation of regulated behavior prompted consideration of how to additionally accommodate sparseness in particular response categories—leading to our construction of a bivariate regime-switching ZI-OU model. Extending prior work, we embedded a Markov-based regime switching into the age-conditional SDE, thus representing (a) transitions in and out of the zero state, (b) the dynamic interdependencies between two components—EP and PR—of the self-regulation processes, and (c) the age differences in those transitions and interdependencies. Estimation was accomplished in a Bayesian framework using MCMC algorithms that were found to perform well in a range of potential data situations. The proposed model opens up possibility for re-examination of many existing data sets—and opportunity to extract knowledge about development of many dynamic processes. We are excited about the push into regime-switching SDEs and the new affordances they provide for testing and revising theoretical propositions about dynamic and change processes.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: JAGS Code

JAGS selects an appropriate MCMC sampler for the parameters automatically from a list of possible samplers, including Gibbs sampler (Geman & Geman, 1984), slice sampler (Neal, 2003), Metropolis–Hastings (Hastings, 1970), and several other possibilities.

```

model{
  for (i in 1:N) {
    # Specify the probabilities in the Latent Regime Transition Model
    # The first process
    MIPMix[1, i, 1, 1] ← 0.5
    MIPMix[1, i, 1, 2] ← 0.5
    for (j in 2:EulerSubLenLastObs[i]) {
      MIPMix[1, i, j, 1] ← exp(Ebeta[(midx[1, i, j - 1] -
        1) * 2 + midx[2, i, j - 1], 1, MixLogiCovar[i,2] + 1])
      MIPMix[1, i, j, 2] ← exp(Ebeta[(midx[1, i, j - 1] -
        1) * 2 + midx[2, i, j - 1], 2, MixLogiCovar[i,2] + 1])
    }

    # The second process
    MIPMix[2, i, 1, 1] ← 0.5
    MIPMix[2, i, 1, 2] ← 0.5
    for (j in 2:EulerSubLenLastObs[i]) {
      MIPMix[2, i, j, 1] ← exp(Pbeta[(midx[1, i, j - 1] -
        1) * 2 + midx[2, i, j - 1], 1, MixLogiCovar[i,2] + 1])
      MIPMix[2, i, j, 2] ← exp(Pbeta[(midx[1, i, j - 1] -
        1) * 2 + midx[2, i, j - 1], 2, MixLogiCovar[i,2] + 1])
    }

    # Specify the regime-switching OU model at time point 1
    for (j in 1:NZ) {
      Y[j, 1, i] ~ dnorm(0.00000E+00, 0.01)
    }
    midx[1, i, 1] ~ dcat(MIPMix[1, i, 1, 1:NMixComp])
  }
}

```

```

midx[2, i, 1] ~ dcat(MIPMix[2, i, 1, 1:NMixComp])
for (j in 2:EulerSubLenLastObs[i]) {

  # Specify the Latent Regime Transition Model
  midx[1, i, j] ~ dcat(MIPMix[1, i, j, 1:NMixComp])
  midx[2, i, j] ~ dcat(MIPMix[2, i, j, 1:NMixComp])

  # Specify the regime-switching OU Model

  Y[1, j, i] ~ dnorm(y $\mu$ [1, j, i], tau[1, j, i])
  y $\mu$ [1, j, i]  $\leftarrow$  Y[1, j - 1, i] + dt1[j - 1, i] *
    dynpar[midx[1, i, j - 1], 1] * (dynpar[midx[1,
      i, j - 1], 2] - Y[1, j - 1, i])
  tau[1, j, i]  $\leftarrow$  1/(dt1[j - 1, i] * dynpar[midx[1,
    i, j - 1], 3])
  Y[2, j, i] ~ dnorm(y $\mu$ [2, j, i], tau[2, j, i])
  y $\mu$ [2, j, i]  $\leftarrow$  Y[2, j - 1, i] + dt1[j - 1, i] *
    dynpar[midx[2, i, j - 1], 4] * (dynpar[midx[2,
      i, j - 1], 5] - Y[2, j - 1, i])
  tau[2, j, i]  $\leftarrow$  1/(dt1[j - 1, i] * dynpar[midx[2,
    i, j - 1], 6])
}
for (j in (EulerSubLenLastObs[i] + 1):EulerLenLastObsA1) {
  midx[1, i, j]  $\leftarrow$  1
  midx[2, i, j]  $\leftarrow$  1
}
}

# Prior of parameters in OU model
for (k in 1:1) {
  dynpar[k, 1]  $\leftarrow$  60
  dynpar[k, 2]  $\leftarrow$  0.00000E+00
  dynpar[k, 3]  $\leftarrow$  1.00000E-04
  dynpar[k, 4]  $\leftarrow$  60
  dynpar[k, 5]  $\leftarrow$  0.00000E+00
  dynpar[k, 6]  $\leftarrow$  1.00000E-04
}
for (k in 2:2) {
  dynpar[k, 1] ~ dnorm(0.2, 0.01) T(0.00000E+00, )
  dynpar[k, 2] ~ dnorm(2, 0.01)
  dynpar[k, 3] ~ dgamma(2, 0.01)
  dynpar[k, 4] ~ dnorm(0.2, 0.01) T(0.00000E+00, )
  dynpar[k, 5] ~ dnorm(2, 0.01)
  dynpar[k, 6] ~ dgamma(2, 0.01)
}

# Prior of parameters in the Latent Regime Transition Model
# parameters for EP in matrix form
Ebeta[1, 1, 1]  $\leftarrow$  Ebetav[1]
Ebeta[1, 2, 1]  $\leftarrow$  0.00000E+00

```

```

Ebeta[2, 1, 1] ← Ebetav[2]
Ebeta[2, 2, 1] ← 0.00000E+00
Ebeta[3, 1, 1] ← 0.00000E+00
Ebeta[3, 2, 1] ← Ebetav[3]
Ebeta[4, 1, 1] ← 0.00000E+00
Ebeta[4, 2, 1] ← Ebetav[4]

# free parameters for EP
Ebetav[1] ~ dnorm(0.00000E+00, 1.00000E-04)
Ebetav[2] ~ dnorm(0.00000E+00, 1.00000E-04)
Ebetav[3] ~ dnorm(0.00000E+00, 1.00000E-04)
Ebetav[4] ~ dnorm(0.00000E+00, 1.00000E-04)

# parameters for PR in matrix form
Pbeta[1, 1, 1] ← Pbetav[1]
Pbeta[1, 2, 1] ← 0.00000E+00
Pbeta[2, 1, 1] ← 0.00000E+00
Pbeta[2, 2, 1] ← Pbetav[2]
Pbeta[3, 1, 1] ← Pbetav[1]
Pbeta[3, 2, 1] ← 0.00000E+00
Pbeta[4, 1, 1] ← 0.00000E+00
Pbeta[4, 2, 1] ← Pbetav[2]

# free parameters for PR
Pbetav[1] ~ dnorm(0.00000E+00, 1.00000E-04)
Pbetav[2] ~ dnorm(0.00000E+00, 1.00000E-04)

# parameters at 48 months
Ebeta[1, 1, 2] ← Ebetav[1]
Ebeta[1, 2, 2] ← 0.00000E+00
Ebeta[2, 1, 2] ← Ebetav[2]
Ebeta[2, 2, 2] ← 0.00000E+00
Ebeta[3, 1, 2] ← 0.00000E+00
Ebeta[3, 2, 2] ← Ebetav[3] + Egammav[1]
Ebeta[4, 1, 2] ← 0.00000E+00
Ebeta[4, 2, 2] ← Ebetav[4]
Egammav[1] ~ dnorm(0.00000E+00, 1.00000E-04)
Pbeta[1, 1, 2] ← Pbetav[1] + Pgammav[1]
Pbeta[1, 2, 2] ← 0.00000E+00
Pbeta[2, 1, 2] ← 0.00000E+00
Pbeta[2, 2, 2] ← Pbetav[2]
Pbeta[3, 1, 2] ← Pbetav[1] + Pgammav[2]
Pbeta[3, 2, 2] ← 0.00000E+00
Pbeta[4, 1, 2] ← 0.00000E+00
Pbeta[4, 2, 2] ← Pbetav[2]
Pgammav[1] ~ dnorm(0.00000E+00, 1.00000E-04)
Pgammav[2] ~ dnorm(0.00000E+00, 1.00000E-04)
}

```

Appendix B: Full Conditional Distributions of Parameters

Let $\mathbf{x}_i = (\mathbf{x}_{i,t_{i,1}}, \dots, \mathbf{x}_{i,t_{i,T_i}})$, $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\mathbf{l}_i = (\mathbf{l}_{i,1}, \dots, \mathbf{l}_{i,T_i})$, $\mathbf{l} = (\mathbf{l}_1, \dots, \mathbf{l}_n)$, $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_R\}$. Let $\boldsymbol{\alpha}$ consist of the free parameters in all $\boldsymbol{\alpha}_{r_1 s_1 s_2}^{\text{EP}}$ and $\boldsymbol{\alpha}_{r_2 s_1 s_2}^{\text{PR}}$ in Eq. (8). The posterior distribution of $\boldsymbol{\theta}$, $\boldsymbol{\alpha}$, and \mathbf{l} is

$$P(\mathbf{l}, \boldsymbol{\theta}, \boldsymbol{\alpha} | \mathbf{x}) \propto P(\mathbf{x} | \mathbf{l}, \boldsymbol{\theta}) P(\mathbf{l} | \boldsymbol{\alpha}) p(\boldsymbol{\theta}) p(\boldsymbol{\alpha}).$$

The distribution of the process under the Euler approximation and given the latent regime indicator and the parameters, $P(\mathbf{x} | \mathbf{l}, \boldsymbol{\theta})$, is

$$\begin{aligned} P(\mathbf{x} | \mathbf{l}, \boldsymbol{\theta}) &= \prod_{i=1}^n \prod_{k=0}^{T_n-1} P(\mathbf{x}_{i,t_{i,k+1}} | \mathbf{x}_{i,t_{i,k}}, \mathbf{l}, \boldsymbol{\theta}) \\ &= \prod_{i=1}^n \prod_{k=0}^{T_n-1} (2\pi)^{-q/2} |\Delta t_{i,k} \mathbf{S}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}}) \mathbf{S}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}})^T|^{-1/2} \\ &\quad \exp \left\{ -\frac{1}{2} \left(\Delta \mathbf{x}_{i,t_{i,k}} - \mathbf{f}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}}) \Delta t_{i,k} \right)^T \left(\Delta t_{i,k} \mathbf{S}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}}) \mathbf{S}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}})^T \right)^{-1} \right. \\ &\quad \left. \left(\Delta \mathbf{x}_{i,t_{i,k}} - \mathbf{f}(\mathbf{x}_{i,t_{i,k}}, \boldsymbol{\theta}_{\mathbf{l}_{i,k}}) \Delta t_{i,k} \right) \right\}. \end{aligned} \quad (17)$$

The distribution of the latent regime indicator given the parameters, $P(\mathbf{l} | \boldsymbol{\alpha})$, is

$$\begin{aligned} P(\mathbf{l} | \boldsymbol{\alpha}) &= \prod_{i=1}^n \prod_{k=0}^{T_n-1} \prod_{j=1}^q p(l_{j t_{i,k+1}} | \mathbf{l}_{i,k}) \\ &= \prod_{i=1}^n \prod_{k=0}^{T_n-1} \prod_{j=1}^q \frac{\exp \{ \boldsymbol{\alpha}_{l_{j t_{i,k+1}} \mathbf{l}_{i,k}}^{(j)T} \mathbf{u}_i \}}{\sum_{r=1}^R \exp \{ \boldsymbol{\alpha}_{r, \mathbf{l}_{i,k}}^{(j)T} \mathbf{u}_i \}}. \end{aligned} \quad (18)$$

The full conditional distribution of $l_{i,k}^{\text{EP}}$ is

$$P(l_{i,k}^{\text{EP}} | \cdot) \propto P(\text{EP}_{i,t_{i,k+1}} | \text{EP}_{i,t_{i,k}}, l_{i,k}^{\text{EP}}, \boldsymbol{\theta}) P(l_{i,k+1}^{\text{EP}} | \mathbf{l}_{i,k}, \boldsymbol{\alpha}) P(l_{i,k}^{\text{EP}} | \mathbf{l}_{i,k-1}, \boldsymbol{\alpha}), \quad (19)$$

which is a categorical distribution. The probabilities of all categories can be calculated by standardizing the right-hand side of Eq. (19).

The full conditional distribution of $\boldsymbol{\alpha}$ is proportional to $P(\mathbf{l} | \boldsymbol{\alpha}) P(\boldsymbol{\alpha})$, where $P(\boldsymbol{\alpha})$ is the prior distribution of $\boldsymbol{\alpha}$ defined in Eq. (8). The full conditional distribution of $\boldsymbol{\alpha}$ is not standard. MCMC algorithms may be employed to generate samples from the distribution, for instance, via Metropolis–Hastings algorithm (Hastings, 1970) or the slice sampler (Neal, 2003).

The full conditional distribution of $\boldsymbol{\theta}$ is model-dependent. In the ZI-OU model, given the prior distributions in Eq. (7), the full conditional distributions of the parameters in the OU model are as follows.

$$\begin{aligned} \beta_{\text{EP},l} | \cdot &\sim N(\tilde{\beta}_{\text{EP},0l}, \tilde{\sigma}_{\beta_{\text{EP},l}}^2) I(0, \infty), & \beta_{\text{PR},l} | \cdot &\sim N(\tilde{\beta}_{\text{PR},0l}, \tilde{\sigma}_{\beta_{\text{PR},l}}^2) I(0, \infty), \\ \mu_{\text{EP},l} | \cdot &\sim N(\tilde{\mu}_{\text{EP},0l}, \tilde{\sigma}_{\mu_{\text{EP},l}}^2), & \mu_{\text{PR},l} | \cdot &\sim N(\tilde{\mu}_{\text{PR},0l}, \tilde{\sigma}_{\mu_{\text{PR},l}}^2), \end{aligned}$$

$$\sigma_{EP,l}|\cdot \sim IG(\tilde{a}_{EP,l}, \tilde{a}_{EP,2l}), \quad \sigma_{PR,l}|\cdot \sim IG(\tilde{a}_{PR,l}, \tilde{a}_{PR,2l}),$$

where

$$\begin{aligned} \tilde{\sigma}_{\beta_{EP,l}}^2 &= \left(\frac{1}{\sigma^2} \sum_{i=1}^n \sum_{l_{i,k}^{EP}=l} (\mu_{EP,l_{i,k}^{EP}} - EP_{i,t_{i,k}})^2 \Delta t_{i,k} + \frac{1}{\sigma_{\beta_{EP,l}}^2} \right)^{-1}, \\ \tilde{\beta}_{EP,0l} &= \tilde{\sigma}_{\beta_{EP,l}}^2 \left(\frac{1}{\sigma^2} \sum_{i=1}^n \sum_{l_{i,k}^{EP}=l} (\mu_{EP,l_{i,k}^{EP}} - EP_{i,t_{i,k}})(EP_{i,t_{i,k+1}} - EP_{i,t_{i,k}}) + \frac{\beta_{EP,0l}}{\sigma_{\beta_{EP,l}}^2} \right), \\ \tilde{\sigma}_{\mu_{EP,l}}^2 &= \left(\frac{\beta_{EP,l}^2}{\sigma^2} \sum_{i=1}^n \sum_{l_{i,k}^{EP}=l} \Delta t_{i,k} + \frac{1}{\sigma_{\mu_{EP,l}}^2} \right)^{-1}, \\ \tilde{\mu}_{EP,0l} &= \tilde{\sigma}_{\mu_{EP,l}}^2 \left(\frac{\beta_{EP,l}}{\sigma^2} \sum_{i=1}^n \sum_{l_{i,k}^{EP}=l} (EP_{i,t_{i,k+1}} - EP_{i,t_{i,k}} + \beta_{EP,l} EP_{i,t_{i,k}} \Delta t_{i,k}) + \frac{\mu_{EP,0l}}{\sigma_{\mu_{EP,l}}^2} \right), \\ \tilde{a}_{EP,1l} &= a_{EP,1l} + \frac{1}{2} \sum_{i=1}^n \sum_{k=0}^{T_i-1} I(l_{i,k}^{EP} = l), \text{ and} \\ \tilde{a}_{EP,2l} &= a_{EP,2l} + \frac{1}{2} \sum_{i=1}^n \sum_{l_{i,k}^{EP}=l} \frac{((EP_{i,t_{i,k+1}} - EP_{i,t_{i,k}}) - \beta_{EP,l_{i,k}^{EP}} (\mu_{EP,l_{i,k}^{EP}} - EP_{i,t_{i,k}}) \Delta t_{i,k})^2}{\Delta t_{i,k}}, \end{aligned}$$

and the parameters for the PR process can be derived similarly.

References

- Ait-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics*, 36(2), 906–937.
- Arminger, G. (1986). Linear stochastic differential equation models for panel data with unobserved variables. In N. Tuma (Ed.), *Sociological methodology 1986* (pp. 187–212). San Francisco: Jossey-Bass.
- Bai, Y., & Wu, L. (2018). Analytic value function for optimal regime-switching pairs trading rules. *Quantitative Finance*, 18(4), 637–654. <https://doi.org/10.1080/14697688.2017.1336281>.
- Baumeister, R. F., & Vohs, K. D. (2007). Self-regulation, ego depletion, and motivation. *Social and Personality Psychology Compass*, 1, 115–128. <https://doi.org/10.1111/j.1751-9004.2007.00001.x>.
- Beaulieu, J. M., Jhwueng, D. C., Boettiger, C., & O'Meara, B. C. (2012). Modeling stabilizing selection: Expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution*, 66(8), 2369–2383. <https://doi.org/10.1111/j.1558-5646.2012.01619.x>.
- Beskos, A., Papaspiliopoulos, O., & Roberts, G. (2009). Monte Carlo maximum likelihood estimation for discretely observed diffusion processes. *The Annals of Statistics*, 37(1), 223–245.
- Beskos, A., Papaspiliopoulos, O., Roberts, G., & Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3), 333–382. (with Discussion).
- Buss, K. A., & Goldsmith, H. H. (1998). Fear and anger regulation in infancy: Effects on the temporal dynamics of affective expression. *Child Development*, 69(2), 359–374. <https://doi.org/10.1111/j.1467-8624.1998.tb06195.x>.
- Calvet, L. E., & Fisher, A. J. (2004). How to forecast long-run volatility: Regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2(1), 49–83. <https://doi.org/10.1093/jfinec/nbh003>.
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press.
- Chow, S.-M., Bendežú, J. J., Cole, P. M., & Ram, N. (2016). A comparison of two-stage approaches for fitting nonlinear ordinary differential equation models with mixed effects. *Multivariate Behavioral Research*, 51(2–3), 154–184.
- Chow, S.-M., Grimm, K. J., Guillaume, F., Dolan, C. V., & McArdle, J. J. (2013). Regime-switching bivariate dual change score model. *Multivariate Behavioral Research*, 48(4), 463–502.
- Chow, S.-M., Lu, Z., Sherwood, A., & Zhu, H. (2016). Fitting nonlinear ordinary differential equation models with random effects and unknown initial conditions using the stochastic approximation expectation–maximization (SAEM) algorithm. *Psychometrika*, 81(1), 102–134. <https://doi.org/10.1007/s11336-014-9431-z>.

- Chow, S.-M., Witkiewitz, K., Grasman, R. P. P., & Maisto, S. A. (2015). The cusp catastrophe model as cross-sectional and longitudinal mixture structural equation models. *Psychological Methods, 20*, 142–164.
- Chow, S.-M., & Zhang, G. (2013). Nonlinear regime-switching state-space (RSSS) models. *Psychometrika, 78*(4), 740–768.
- Cole, P. M., BendeZú, J. J., Ram, N., & Chow, S.-M. (2017). Dynamical systems modeling of early childhood self-regulation. *Emotion, 17*(4), 684–699. <https://doi.org/10.1037/emo0000268>.
- Cole, P. M., Tan, P. Z., Hall, S. E., Zhang, Y., Crnic, K. A., Blair, C. B., et al. (2011). Developmental changes in anger expression and attention focus: Learning to wait. *Developmental Psychology, 47*(4), 1078.
- Coleman, J. S. (1968). The mathematical study of change. In H. M. Blalock Jr. & A. Blalock (Eds.), *Methodology in social research* (pp. 428–478). New York: McGraw-Hill.
- Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research, 28*, 131–157.
- Dolan, C. V., Schmittmann, V. D., Lubke, G. H., & Neale, M. C. (2005). Regime switching in the latent growth curve mixture model. *Structural Equation Modeling, 12*(1), 94–119.
- Durham, G. B., & Gallant, A. R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business & Economic Statistics, 20*(3), 297–316.
- Elerian, O., Chib, S., & Shephard, N. (2001). Likelihood inference for discretely observed nonlinear diffusions. *Econometrica, 69*(4), 959–993.
- Elliott, R. J., Aggoun, L., & Moore, J. (1995). *Hidden Markov models: Estimation and control*. New York: Springer.
- Fox, E. B., Sudderth, E. B., Jordan, M. I., & Willsky, A. S. (2010). Bayesian nonparametric methods for learning Markov switching processes. *IEEE Signal Processing Magazine, 27*(6), 43–54. <https://doi.org/10.1109/MSP.2010.937999>.
- Fraley, C., Raftery, A. E., Murphy, T. B., & Scrucca, L. (2012). Mclust version 4 for R: Normal mixture modeling for model-based clustering, classification, and density estimation (No. 597). Department of Statistics, University of Washington.
- Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *Neuroimage, 63*, 310–319. <https://doi.org/10.1093/jfnec/nbh003>.
- Gelman, A. (1996). Inference and monitoring convergence. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov chain Monte Carlo in practice* (pp. 131–143). Boca Raton, FL: CRC Press LLC.
- Gelman, A., Meng, X.-L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica, 6*(4), 733–760.
- Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 721–741.
- Ghysels, E., McCulloch, R. E., & Tsay, R. S. (1998). Bayesian inference for periodic regime-switching models. *Journal of Applied Econometrics, 13*(2), 129–143.
- Gilks, W. R., Best, N. G., & Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society Series C (Applied Statistics), 44*(4), 455–472.
- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*. Boca Raton, FL: CRC Press.
- Goldsmith, H. H., & Reilly, J. (1993). *Laboratory assessment of temperament-preschool version*. University of Oregon: Unpublished manual.
- Golightly, A., & Wilkinson, D. J. (2008). Bayesian inference for nonlinear multivariate diffusion models observed with error. *Computational Statistics & Data Analysis, 52*(3), 1674–1693. <https://doi.org/10.1016/j.csda.2007.05.019>.
- Hall, D. B. (2000). Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics, 56*(4), 1030–1039.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their application. *Biometrika, 57*, 97–100.
- Jones, R. H. (1984). Fitting multivariate models to unequally spaced data. In E. Parzen (Ed.), *Time series analysis of irregularly observed data* (Vol. 25, pp. 158–188). New York: Springer.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.
- Kim, C.-J., & Kim, J. (2015). Bayesian inference in regime-switching ARMA models with absorbing states: the dynamics of the ex-ante real interest rate under regime shifts. *Journal of Business & Economic Statistics, 33*(4), 566–578. <https://doi.org/10.1080/07350015.2014.979995>.
- Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.
- Kloeden, P. E., & Platen, E. (1999). *Numerical solution of stochastic differential equations*. Berlin: Springer.
- Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychology, 18*, 199–214.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics, 34*(1), 1–14. <https://doi.org/10.2307/1269547>.
- Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology, 44*(2), 446–456.
- Lindström, E. (2012). A regularized bridge sampler for sparsely sampled diffusions. *Statistics and Computing, 22*(2), 615–623. <https://doi.org/10.1007/s11222-011-9255-y>.
- Lu, Z.-H., Chow, S.-M., Sherwood, A., & Zhu, H. (2015). Bayesian analysis of ambulatory blood pressure dynamics with application to irregularly spaced sparse data. *The Annals of Applied Statistics, 9*(3), 1601–1620. <https://doi.org/10.1214/15-AOAS846>.

- Maisto, S. A., Xie, F. C., Witkiewitz, K., Ewart, C. K., Connors, G. J., Zhu, H., et al. (2017). How chronic self-regulatory stress, poor anger regulation, and momentary affect undermine treatment for alcohol use disorder: Integrating social action theory and the dynamic model of relapse. *Journal of Social and Clinical Psychology, 36*, 238–263. <https://doi.org/10.1521/jscp.2017.36.3.238>.
- Mbalawata, I. S., Särkkä, S., & Haario, H. (2013). Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering. *Computational Statistics, 28*(3), 1195–1223. <https://doi.org/10.1007/s00180-012-0352-y>.
- Molenaar, P. C. M., & Newell, K. M. (2003). Direct fit of a theoretical model of phase transition in oscillatory finger motions. *British Journal of Mathematical and Statistical Psychology, 56*, 199–214. <https://doi.org/10.1348/000711003770480002>.
- Neal, R. M. (2003). Slice sampling. *The Annals of Statistics, 31*(3), 705–767. <https://doi.org/10.1214/aos/1056562461>.
- Nylund, K. L., Muthén, B., Nishina, A., Bellmore, A., & Graham, S. (2006). Stability and instability of peer victimization during middle school: Using latent transition analysis with covariates, distal outcomes, and modeling extensions. https://scholar.google.com/citations?user=cKtI1DQAAAAJ&hl=en#d=gs_md_cita-d&u=%2Fcitations%3Fview_op%3Dview_citation%26hl%3Den%26user%3DcKtI1DQAAAAJ%26citation_for_view%3DcKtI1DQAAAAJ%3AjjCSPb-OGe4C%26tzm%3D360.
- Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods, 16*, 468–490. <https://doi.org/10.1037/a0024375>.
- Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2016). Bayesian data analysis with the bivariate hierarchical Ornstein–Uhlenbeck process model. *Multivariate Behavioral Research, 51*(1), 106–119. <https://doi.org/10.1080/00273171.2015.1110512>.
- Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*(2), 199–215.
- Oud, J.H.L., & Singer, H. (Eds.). (2008). *Special issue: Continuous time modeling of panel data* (Vol. 62(1)).
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125). Wien: Technische Universit.
- Ramsay, J. O., Hooker, G., Campbell, D., & Cao, J. (2007). Parameter estimation for differential equations: A generalized smoothing approach. *Journal of Royal Statistical Society: Series B, 69*(5), 741–796. <https://doi.org/10.1111/j.1467-9868.2007.00610.x>. (with Discussion).
- Roberts, G. O., & Stramer, O. (2001). On inference for partial observed nonlinear diffusion models using the Metropolis–Hastings algorithm. *Biometrika, 88*, 603–621. <https://doi.org/10.1093/biomet/88.3.603>.
- Roeder, K., Lynch, K. G., & Nagin, D. S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association, 94*, 766–776. <https://doi.org/10.1080/0022250x.2010.509524>.
- Särkkä, S. (2013). *Bayesian filtering and smoothing*. Hillsdale, NJ: Cambridge University.
- Singer, H. (1992). The aliasing-phenomenon in visual terms. *Journal of Mathematical Sociology, 14*(1), 39–49.
- Singer, H. (2010). SEM modeling with singular moment matrices. Part I: ML-estimation of time series. *The Journal of Mathematical Sociology, 34*(4), 301–320. <https://doi.org/10.1080/0022250x.2010.532259>.
- Singer, H. (2012). SEM modeling with singular moment matrices. Part II: ML-estimation of sampled stochastic differential equations. *The Journal of Mathematical Sociology, 36*(1), 22–43. <https://doi.org/10.1080/0022250x.2010.532259>.
- Solo, V. (2002). Identification of a noisy stochastic heat equation with the EM algorithm. In *Proceedings of the 41st IEEE conference on decision and control, 2002*. (Vol. 4, pp. 4505–4508). <https://doi.org/10.1109/CDC.2002.1185083>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583–639.
- Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical review, 36*(5), 823.
- Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. arXiv (preprint).
- Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods, 17*, 176–192. <https://doi.org/10.1037/a0027543>.
- Wilhelm, F. H., Grossman, P., & Muller, M. I. (2012). Bridging the gap between the laboratory and the real world: Integrative ambulatory psychophysiology. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 210–234). New York: Guilford.
- Yang, J.-W., Tsai, S.-Y., Shyu, S.-D., & Chang, C.-C. (2016). Pairs trading: The performance of a stochastic spread model with regime switching-evidence from the S&P 500. *International Review of Economics & Finance, 43*, 139–150. <https://doi.org/10.1016/j.iref.2015.10.036>.
- Yang, M., & Chow, S.-M. (2010). Using state-space model with regime switching to represent the dynamics of facial electromyography (EMG) data. *Psychometrika: Application and Case Studies, 74*(4), 744–771.
- Yümlü, M. S., Gürgen, F. S., Cemgil, A. T., & Okay, N. (2015). Bayesian changepoint and time-varying parameter learning in regime switching volatility models. *Digital Signal Processing, 40*, 198–212. <https://doi.org/10.1016/j.dsp.2015.02.001>.